Isolated Polynucleotides Encoding DsrA, A Protein Conferring Serum Resistance To *H. ducreyi*, And Methods And Compositions Comprising The Same

Statement of Federal Support

This invention was made with United States Government support under grant numbers AI 40263 and A126837 from the National Institutes of Health (Public Health Service). The United States Government has certain rights to this invention.

Field of The Invention

This invention relates to proteins that are involved in the serum resistance of *H. ducreyi*.

Background of the Invention

Haemophilus ducreyi is the etiologic agent of chancroid, a genital ulcer disease transmitted by sexual contact. *See, e.g.*, Albritton, W. L., *Microbiol Rev.* **53**:377-89 (1989); Trees, D. L., and S. A. Morse, *Clin Microbiol Rev.* **8**, 357-375 (1995). Chancroid has gained importance recently because it has been implicated as an independent risk factor for the heterosexual transmission of HIV in Africa. *See* Albritton, *supra*, Trees, *supra*; R.M. Greenblattet et al., *AIDS* **2,** 47-50 (1988); Jessamine, P. G., and A. R. Ronald, *Med Clin North Am.* **74**, 1417-31 (1990); Plummer, F. A. et al., *J Infect Dis.* **161**, 810-1 (1990); D. L., and S. A. Morse, *Clin. Microbiol Rev.* **8**, 357-375 (1995); Wasserheit, J. N., *Sex Trans Dis.* **19**, 61-77 (1991).

Serum resistance has been shown in numerous bacterial systems to be critical for the survival of invading bacterial and the establishment of disease, since

mutations which result in the loss of serum resistance renders several bacterial pathogens avirulent. *See. e.g.*, Blaser, M. J., *American Journal of the Medical Sciences*. **306**, 325-9 (1993); Corbeil, L. B., *Canadian Journal of Veterinary Research*. **54**,S57-62 (1990), Mobley, H. L. et al., *Kidney International - Supplement*. **47**, S129-36 (1994); Rice, P. A., *Clinical Microbiology Reviews*. **2**, S112-7 (1989); and Stull, T. L., and J. J. LiPuma, *Medical Clinics of North America*. **75**, 287-9 (1991). In most systems, the serum resistance phenotype is the product of multiple genes. *H. ducreyi* is resistant to high levels of normal human serum (NHS; up to 50%). Early studies on *H. ducreyi* serum resistance by Odumeru and colleagues concluded that truncation of LOS in several strains was associated with avirulence and loss of serum resistance (*see* Odumeru, J. A., *et al.*, *Infect. Immun.* **43,** 607-611 (1984); Odumeru, J. A. et al., *Infect. Immun.* **50**, 495-9 (1985); Odumeru, J. A. et al., *J Med Microbiol*. **23,** 155-62 (1987)), whereas a recent study came to the opposite conclusion. *See* Hiltke, T. J. *et al.*, *Microb Path.* **26**,93-102 (1999)

Originally described as a cell spreading factor, vitronectin is now recognized as a multifunctional regulatory adhesive glycoprotein involved in a variety of extracellular processes such as the attachment and spreading of normal and neoplastic cells, as well as the function of the complement and coagulation pathways. Integrins are transmembrane $\alpha\beta$ heterodimer receptors expressed on a wide variety of cells which are involved in extracellular matrix interactions. The ligands for several of the integrins are adhesive extracellular matrix (ECM) proteins such as fibronectin, vitronectin, collagens and laminin.

Proteins or fragments thereof that are able to interfere with vitronectin binding to various integrins and to block integrin-mediated cell attachment to extracellular matrix proteins are useful in preventing the attachment of the bacteria to the host organism, and thus infection of the host.

The ability to use a protein or antibody that interferes with vitronectin binding in a vaccine against *H ducreyi* is desirable. These kinds of proteins are believed to be highly conserved among strains of a particular type of bacteria in that they are the protein molecules that mediate attachment by bonding bacteria to host cells, the initial step in the infection process. A vaccine against *H. ducreyi* comprising a protein or antibody that would interfere with vitronectin binding

would be effective against a broad array of types and strains of *H. ducreyi*. The use of such a vaccine may prevent adherence of the bacteria to the tissue of the host animal. In that adherence is one of the initial step in *H. ducreyi* infection, accordingly, preventing or limiting the infection at this point would be advantageous.

In view of the foregoing, it would be desirable to determine the mechanism of serum resistance in *H. ducreyi*. Additionally, the development of an effective vaccine against *H. ducreyi* would be advantageous.

Summary of the Invention

Certain objects, advantages and novel features of the invention will be set forth in the description that follows, and will become apparent to those skilled in the art upon examination of the following, or may be learned with the practice of the invention.

The present invention is based in the inventor's discovery that a protein, referred to herein as DsrA, has been found to play a critical role in the resistance of *H. ducreyi* to normal human serum.

Accordingly, one aspect of the invention is a polynucleotide (e.g., DNA) that encodes the protein DsrA. Particularly preferred is the DNA of **SEQ ID NO:1**, which encodes the protein DsrA set forth in **SEQ ID NO:2**.

An additional aspect of the invention is the isolated protein DsrA, which protein may vary in molecular weight between 28 and 35 kilodaltons, depending on whether the particular DsrA protein sequence comprises one, two or three copies of the amino acid heptamer NTHNINK (**SEQ ID NO:19**).

Expression vectors and host cells expressing DsrA are also an aspect of the invention. Antibodies against DsrA and antisense molecules of DsrA are a further aspect of the present invention.

Vaccines against H. ducreyi comprising proteins, polynucleotides and expression vectors of DsrA are a further aspect of the invention.

Also an aspect of this invention is an isogenic mutant (FX517) of *H. ducreyi* strain 35000 that does not express DsrA, which mutant finds use in an attenuated vaccine against *H. ducreyi*.

The foregoing and other aspects of the present invention are explained in detail in the specification set forth below.

Brief Description of the Drawings

FIG. 1 is a photograph of Western Blot illustrating the distribution of the DsrA protein and summary of serum resistance of *H. ducreyi* strains. Total cellular proteins from geographically diverse *H. ducreyi* strains were subjected to SDS-PAGE and Western blotting using anti-DsrA mouse sera. Bound antibody was detected with alkaline phosphatase-conjugated secondary antibody and BCIP/NBT substrate. An additional twelve *H. ducreyi* strains also expressed a 28-35 kDa protein which reacted with this serum (data not shown). The names of strains are indicated above each lane. Shown to the left of the gel are molecular weight standards, where the abbreviation kDa means kilodaltons. R refers to resistant to 50% NHS; S, sensitive to 50% NHS, an asterisk indicates that resistance to NHS was indeterminate. The data in **FIG. 1** are compiled from experiments done on at least three separate days.

FIG. 2 is a schematic illustration of the restriction map of the *dsrA* region and PCR products thereof. The *dsrA* open reading frame is boxed. The restriction sites are indicated. The numbered arrows indicate direction and position of the *dsrA* oligos used for PCR. The letter KS and T7 (promoter) refer to the vector primers used in the vector-anchored PCR reactions. V-A PCR refers to vector-anchored PCR; P refers to a promoter. The jagged lines represent approximately 2 kb of sequence not shown downstream of the *dsrA* locus.

FIG. 3 sets forth the DNA sequence (**SEQ ID NO:1**) and deduced amino acid sequence (**SEQ ID NO:2**) of the *dsrA* locus. The putative -35 and -10 promoter sequences are indicated and underlined. A putative ribosome binding site is labeled RBS and underlined. Twenty one amino acids comprising the signal peptide are underlined. The stop codon TAA is indicated with an asterisk. The opposing arrows show a potential stem loop transcription terminator.

FIG. 4 sets forth a comparison of the amino acid sequence of DsrA (**SEQ ID NO:2**) with the UspA2 protein of *M. catarrahalis* (**SEQ ID NO:20**) and the YadA protein of *Y. enterocolitica* (**SEQ ID NO:21**). Shaded, boxed residues indicate homologous sequences.

FIG. 5 is a photograph of a SDS-PAGE/Western blot of parent strain 35000 and *dsrA* mutant FX517. Outer membranes were prepared, solubilized at 37°C or 100°C and subjected to SDS-PAGE and Coomassie staining (Panel A). For the Western blot (panel B), outer membranes were solubilized at 100°C, transferred to nitrocellulose and probed with anti-DsrA mouse serum. Bound antibody was detected with alkaline phosphatase-conjugated secondary antibody and BCIP/NBT substrate. The asterices indicate the positions of the DsrA protein. STD, molecular weight standards.

FIG. 6 is a graphical illustration of the bactericidal killing of parent strain 35000 compared with the bactericidal killing of the *dsrA* mutant FX517. Bactericidal killing was performed as described in **FIG. 1**, except that two serum concentrations were utilized. The data presented in **FIGS. 1** and **6** for 35000 with 50% sera are the same data. The data presented for 35000 were obtained in parallel experiments with FX517.

FIG. 7 is a photograph of a SDS-PAGE/Western Blot illustrating Complementation of *dsrA* mutants. Total cellular proteins from the indicated *H. ducreyi* strains were subjected to SDS-PAGE (12%) and Western blotting using anti-DsrA antisera. Bound antibody was detected with horseradish peroxidase-conjugated secondary antibody followed by chemiluminescence. "N" indicates no plasmid present; "+" indicates pUNCH 1260 (*i.e.*, contains the entire *dsrA* ORF from strain 35000 and its putative native promoter as illustrated in **Fig. 2**); "-"indicates pLSKS a vector without insert. Below each strain are shown the summary of bactericidal killing of the complemented *dsrA* mutants. Bactericidal killing was performed as in **FIG. 1** (50% serum), except that the medium used contained streptomycin.

FIG. 8 is a photograph of an SDS-PAGE gel illustrating the analysis of LOS as described in Example 16, below. Crude LOS was prepared as described in the text and subjected to SDS-PAGE and silver staining.

FIG. 9 illustrates a comparison of the deduced amino acid sequences of *dsrA* from strain 35000 (**SEQ ID NO:2**) and eight additional *H. ducreyi* strains (CIP A75, **SEQ ID NO:4**; CIP A77, **SEQ ID NO:6**; CIP542 (CAN), **SEQ ID NO:8**; CIP542 (CDC), **SEQ ID NO:10**; CHIA, **SEQ ID NO:12**; V-1157, **SEQ ID**

NO:14; M90-02, **SEQ ID NO:16** and 406, **SEQ ID NO:18**). Variable regions 1 and 2 are indicated.

FIG. 10 illustrates the promoter regions of *dsrA* from various strains of *H. ducreyi* (35000, CIP542 (CAN), CIP542 (CDC), CHIA, V-1157, M90-02 and 406, **SEQ ID NO:22**; CIP A75 and CIP A77, **SEQ ID NO:23**) and the mutations in the strains CIP A75 and CIP A77, which do not express DsrA. The 5 base-pair deletions present in strains CIP A75 and CIP A77 are shown as hyphens.

FIG. 11 is a graphical illustration showing that efficient attachment of *H. ducreyi* to a keratinocyte cell line requires DsrA expression. *H. ducreyi* were added to HaCaT cells at a MOI of between 1-5:1 and incubated for two hours. After removal of unbound bacteria by extensive washing, CFUs were determined by plating the disrupted monolayer. The data shown in **FIG. 11** are taken from four experiments.

FIG. 12 is an autoradiograph of an SDS-PAGE illustrating the affinity purification of DsrA from whole cells using biotinylated vitronectins (Vn). Biotinylated vitronectins were mixed with surface-iodinated *H. ducreyi* and allowed to bind. After washing unbound vitronectin by centrifugation and washing, *H. ducreyi* were solubilized with a gentle detergent. Total soluble *H. ducreyi* proteins were bound to solid-phase streptavidin-agarose. After washing the streptavidin agarose, bound proteins were eluted by boiling in sample buffer and analysis by SDS-PAGE and autoradiography.

Detailed Description of the Preferred Embodiments

The present invention will now be described more fully hereinafter with reference to the accompanying figures, in which preferred embodiments of the invention are shown. This invention may, however, be embodied in different forms and should not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art

to which this invention belongs. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety.

Amino acid sequences disclosed herein are presented in the amino to carboxy direction, from left to right. The amino and carboxy groups are not presented in the sequence. Nucleotide sequences are presented herein by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented herein in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission, or (for amino acids) by three letter code, in accordance with 37 CFR §1.822 and established usage. *See, e.g.*, *Patent In User Manual*, 99-102 (Nov. 1990) (U.S. Patent and Trademark Office).

Dsr A is an *H. ducreyi* outer membrane protein required for the expression of serum resistance and is encoded by the gene dsrA, described herein. The isolated *H. ducreyi* protein DsrA, and the isolated polynucleotides that encode the protein, are aspects of the present invention. The DsrA protein in its monomer form varies in molecular weight between 28 and 35kDA between different *H. ducreyi* strains in SDS-PAGE and Western blots. The dsrA locus from several *H. ducreyi* strains was sequenced and the deduced amino acid sequences were greater than 85% identical. The major difference between the different strains is found in the amino acid sequence, in which either one, two or three copies of the amino acid sequence NTHNINK (**SEQ ID NO:19**) in the VR2 region of the protein present; these repeats account for the variability in the monomer form of the DsrA observed in SDS-PAGE. DsrA proteins that contain one, two or three copies of the NTHNINK (**SEQ ID NO:19**) in the VR2 region of the protein, and accordingly having a molecular weight of between 28 and 35 kilodaltons, are all within the scope of the present invention. Additionally, DsrA, as used herein, refers to the amino acid sequences of substantially purified DsrA obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source whether natural, synthetic, semi-synthetic, or recombinant.

As used herein, in this context, the term "isolated" means that the protein is significantly free of other proteins. That is, a composition comprising the isolated protein is between 70% and 94% pure by weight. Preferably, the protein is purified. As used herein, the term "purified" and related terms means that the

protein is at least 95% pure by weight, preferably at least 98% pure by weight, and most preferably at least 99% pure by weight.

An "allele" as used herein, is an alternative form of the polynucleotides (i.e., genes) encoding DsrA. Alleles may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Amino acid sequence," as used herein, refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragment thereof, and to naturally occurring or synthetic molecules. Fragments of DsrA are preferably and retain the biological activity or the immunological activity of DsrA. Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, amino acid sequence, and like terms, are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

"Amplification", as used herein, refers to the production of additional copies of a nucleic acid sequence and is generally carried out using polymerase chain reaction (PCR) technologies well known in the art (Dieffenbach, C. W. and G. S. Dveksler, *PCR Primer, a Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y.(1995)).

As used herein, the term "antibody" refers to intact molecules as well as fragments thereof, such as Fa, F(ab´)2, and Fc, which are capable of binding the DsrA protein or an antigenic or epitopic determinant thereof. Antibodies that bind DsrA polypeptides can be prepared using intact polypeptides or fragments containing small peptides of interest as an immunizing antigen. The polypeptide or oligopeptide may be used to immunize an animal and can be derived from the translation of RNA or synthesized chemically and can be conjugated to a carrier protein, if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin and thyroglobulin, keyhole limpet hemocyanin. The

coupled peptide is then used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

The term "antigenic determinant", as used herein, refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to a given region or three-dimensional structure on the protein; these regions or structures are referred to as antigenic determinants. An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense", as used herein, refers to any composition containing nucleotide sequences which are complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules include peptide nucleic acids and may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and block either transcription or translation. The designation "negative" is sometimes used in reference to the antisense strand, and "positive" is sometimes used in reference to the sense strand.

The terms "complementary" or "complementarity," as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, the sequence "A-G-T" binds to the complementary sequence "T-C-A". Complementarity between two single-stranded molecules may be "partial," in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands.

A "deletion", as used herein, refers to a change in the amino acid or nucleotide sequence and results in the absence of one or more amino acid residues or nucleotides.

The term "derivative", as used herein, refers to the chemical modification of a nucleic acid encoding or complementary to DsrA or the encoded DsrA. Such modifications include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative encodes a polypeptide which retains the biological or immunological function of the natural molecule. A derivative polypeptide is one which is modified by glycosylation, pegylation, or any similar process which retains the biological or immunological function of the polypeptide from which it was derived.

The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (*i.e.*, identity). A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or hybridization probe will compete for and inhibit the binding of a completely homologous sequence to the target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% identity). In the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence.

The term "hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing. The term "hybridization complex", as used herein, refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary G and C bases and between complementary A and T bases; these hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences hydrogen bond in an antiparallel configuration. A hybridization complex may be formed in solution (e.g., C_0t or R_0t

analysis) or between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

An "insertion" or "addition", as used herein, refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, as compared to the naturally occurring molecule.

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. "Fragments" are those nucleic acid sequences which are greater than 60 nucleotides than in length, and most preferably includes fragments that are at least 100 nucleotides or at least 1000 nucleotides, and at least 10,000 nucleotides in length.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20 to 25 nucleotides, which can be used in PCR amplification or a hybridization assay, or a microarray. As used herein, oligonucleotide is substantially equivalent to the terms "amplimers", "primers", "oligomers", and "probes", as commonly defined in the art.

The term "sample", as used herein, is used in its broadest sense. A biological sample suspected of containing nucleic acid encoding DsrA, or fragments thereof, or DsrA itself may comprise a bodily fluid, extract from a cell, chromosome, organelle, or membrane isolated from a cell, a cell, genomic DNA, RNA, or cDNA (in solution or bound to a solid support, a tissue, a tissue print, and the like).

The terms "stringent conditions" or "stringency", as used herein, refer to the conditions for hybridization as defined by the nucleic acid, salt, and temperature. These conditions are well known in the art and may be altered in order to identify or detect identical or related polynucleotide sequences. Numerous equivalent conditions comprising either low or high stringency depend on factors such as the length and nature of the sequence (DNA, RNA, base composition),

nature of the target (DNA, RNA, base composition), milieu (in solution or immobilized on a solid substrate), concentration of salts and other components (e.g., formamide, dextran sulfate and/or polyethylene glycol), and temperature of the reactions (within a range from about 5° C. below the melting temperature of the probe to about 20° C. to 25° C. below the melting temperature). One or more factors may be varied to generate conditions of either low or high stringency different from, but equivalent to, the above listed conditions.

A "substitution", as used herein, refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Transformation", as defined herein, describes a process by which exogenous DNA enters and changes a recipient cell. It may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. Such "transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. They also include cells which transiently express the inserted DNA or RNA for limited periods of time.

Polynucleotides of the present invention include those polynucleotides encoding for proteins homologous to, and having essentially the same biological properties as, the protein DsrA disclosed herein. Particularly preferred is the DNA disclosed herein as **SEQ ID NO:1** and encoding the protein DsrA given herein **SEQ ID NO:2**. This definition of polynucleotides of the present invention is intended to encompass natural allelic sequences thereof. Accordingly, other preferred embodiments of the present invention include the polynucleotides set forth herein as **SEQ ID NO:3**, **SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15,** and **SEQ ID NO:17**, which polynucleotide sequences encode the protein sequences set forth as **SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16,** and **SEQ ID NO:18,** respectively. Isolated DNA or cloned genes of the present invention can be of any species of origin, including

mouse, rat, rabbit, cat, porcine, and human, but are preferably of mammalian origin. Polynucleotides that hybridize to any one of the DNA disclosed herein as **SEQ ID NO:1**, **SEQ ID NO:3**, **SEQ ID NO:5**, **SEQ ID NO:7**, **SEQ ID NO:9**, **SEQ ID NO:11**, **SEQ ID NO:13**, **SEQ ID NO:15**, or **SEQ ID NO:17** (or fragments or derivatives thereof which serve as hybridization probes as discussed below) and which code on expression for a protein of the present invention (*e.g.*, a protein according to **SEQ ID NO:2**, **SEQ ID NO:4**, **SEQ ID NO:6**, **SEQ ID NO:8**, **SEQ ID NO:10**, **SEQ ID NO:12**, **SEQ ID NO:14**, **SEQ ID NO:16**, or **SEQ ID NO:18**) are also an aspect of the invention. Conditions which will permit other polynucleotides that code on expression for a protein of the present invention to hybridize to the any one of DNA of **SEQ ID NO:1**, **SEQ ID NO:3**, **SEQ ID NO:5**, **SEQ ID NO:7**, **SEQ ID NO:9**, **SEQ ID NO:11**, **SEQ ID NO:13**, **SEQ ID NO:15**, or **SEQ ID NO:17** disclosed herein can be determined in accordance with known techniques. For example, hybridization of such sequences may be carried out under conditions of reduced stringency, medium stringency or even stringent conditions (*e.g.*, conditions represented by a wash stringency of 35-40% Formamide with 5x Denhardt's solution, 0.5% SDS and 1x SSPE at 37°C; conditions represented by a wash stringency of 40-45% Formamide with 5x Denhardt's solution, 0.5% SDS, and 1x SSPE at 42°C; and conditions represented by a wash stringency of 50% Formamide with 5x Denhardt's solution, 0.5% SDS and 1x SSPE at 42°C, respectively) to any one of the DNA of **SEQ ID NO:1**, **SEQ ID NO:3**, **SEQ ID NO:5**, **SEQ ID NO:7**, **SEQ ID NO:9**, **SEQ ID NO:11**, **SEQ ID NO:13**, **SEQ ID NO:15**, or **SEQ ID NO:17** disclosed herein in a standard hybridization assay. *See, e.g.*, J. Sambrook et al., *Molecular Cloning, A Laboratory Manual* (2d Ed. 1989) (Cold Spring Harbor Laboratory). In general, sequences which code for proteins of the present invention and which hybridize to any one of the DNA of **SEQ ID NO:1**, **SEQ ID NO:3**, **SEQ ID NO:5**, **SEQ ID NO:7**, **SEQ ID NO:9**, **SEQ ID NO:11**, **SEQ ID NO:13**, **SEQ ID NO:15**, or **SEQ ID NO:17** disclosed herein will be at least 75% homologous, 85% homologous, and even 95% homologous or more with the any one of **SEQ ID NO:1**, **SEQ ID NO:3**, **SEQ ID NO:5**, **SEQ ID NO:7**, **SEQ ID NO:9**, **SEQ ID NO:11**, **SEQ ID NO:13**, **SEQ ID NO:15**, or **SEQ ID NO:17** . Further, polynucleotides that code for proteins of the present invention, or polynucleotides that hybridize to any one of

SEQ ID NO:1, **SEQ ID NO:3**, **SEQ ID NO:5**, **SEQ ID NO:7**, **SEQ ID NO:9**, **SEQ ID NO:11**, **SEQ ID NO:13**, **SEQ ID NO:15**, and **SEQ ID NO:17**, but which differ in codon sequence from any one of **SEQ ID NO:1**, **SEQ ID NO:3**, **SEQ ID NO:5**, **SEQ ID NO:7**, **SEQ ID NO:9**, **SEQ ID NO:11**, **SEQ ID NO:13**, **SEQ ID NO:15**, or **SEQ ID NO:17** due to the degeneracy of the genetic code, are also an aspect of this invention. The degeneracy of the genetic code, which allows different nucleic acid sequences to code for the same protein or peptide, is well known in the literature. See, e.g., U.S. Patent No. 4,757,006 to Toole et al. at Col. 2, Table 1.

Although nucleotide sequences which encode DsrA and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring DsrA under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding DsrA or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding DsrA and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences, or fragments thereof, which encode DsrA and its derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents that are well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding DsrA or any fragment thereof.

Knowledge of the nucleotide sequence as disclosed herein in **SEQ ID NO:1**, **SEQ ID NO:3**, **SEQ ID NO:5**, **SEQ ID NO:7**, **SEQ ID NO:9**, **SEQ ID NO:11**, **SEQ ID NO:13**, **SEQ ID NO:15**, and **SEQ ID NO:17**, can be used to generate hybridization probes which specifically bind to the DNA of the present invention or to mRNA to determine the presence of amplification or overexpression of the proteins of the present invention.

The production of cloned genes, recombinant DNA, vectors, transformed host cells, proteins and protein fragments by genetic engineering is well known. *See, e.g.*, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory (1989)), as well as U.S. Patent No. 4,761,371 to Bell et al. at Col. 6 line 3 to Col. 9 line 65; U.S. Patent No. 4,877,729 to Clark et al. at Col. 4 line 38 to Col. 7 line 6; U.S. Patent No. 4,912,038 to Schilling at Col. 3 line 26 to Col. 14 line 12; and U.S. Patent No. 4,879,224 to Wallner at Col. 6 line 8 to Col. 8 line 59. (Applicant specifically intends that the disclosure of all patent references cited herein be incorporated herein in their entirety by reference).

Methods for DNA sequencing which are well known and generally available in the art may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE® (US Biochemical Corp, Cleveland, Ohio), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, Ill.), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE Amplification System marketed by Gibco/BRL (Gaithersburg, Md.). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, Nev.), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, Mass.) and the ABI Catalyst and 373 and 377 DNA Sequencers (Perkin Elmer).

The nucleic acid sequences encoding DsrA may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, "restriction-site" PCR, uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, G., *PCR Methods Applic.* **2**,318-322 (1993)). In particular, genomic DNA is first amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

A vector, as defined herein, is a replicable DNA construct. Vectors, such as plasmids, are used herein either to amplify DNA encoding the proteins of the present invention or to express the proteins of the present invention. An expression vector is a replicable DNA construct in which a DNA sequence encoding the proteins of the present invention is operably linked to suitable control sequences capable of effecting the expression of proteins of the present invention in a suitable host. The need for such control sequences will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites, and sequences which control the termination of transcription and translation. Amplification vectors do not require expression control domains. All that is needed is the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants.

Vectors, as used herein, include plasmids, viruses (*e.g.*, adenovirus, cytomegalovirus), phage, retroviruses and integratable DNA fragments (i.e., fragments integratable into the host genome by recombination). The vector replicates and functions independently of the host genome, or may, in some instances, integrate into the genome itself. Expression vectors preferably contain a promoter and RNA binding sites which are operably linked to the gene to be expressed and are operable in the host organism.

DNA regions are operably linked or operably associated when they are functionally related to each other. For example, a promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation. Generally, operably linked means contiguous and, in the case of leader sequences, contiguous and in reading phase.

Transformed host cells are cells which have been transformed or transfected with vectors containing DNA coding for proteins of the present invention need not express protein. Suitable host cells include prokaryotes, yeast cells, or higher eukaryotic organism cells. Prokaryote host cells include gram negative or gram positive organisms, for example *Escherichia coli* (*E. coli*) or *Bacilli*. Higher eukaryotic cells include established cell lines of mammalian origin

as described below. Exemplary host cells are *E. coli* W3110 (ATCC 27,325), *E. coli B*, *E. coli* X1776 (ATCC 31,537), E. coli 294 (ATCC 31,446). A broad variety of suitable prokaryotic and microbial vectors are available. *E. coli* is typically transformed using a derivative of the plasmid pBR322. See Bolivar et al., *Gene* **2**, 95 (1977). Promoters most commonly used in recombinant microbial expression vectors include the beta-lactamase (penicillinase) and lactose promoter systems (Chang et al., *Nature* **275**, 615 (1978); and Goeddel et al., *Nature* **281**, 544 (1979), a tryptophan (trp) promoter system (Goeddel et al., *Nucleic Acids Res.* **8**, 4057 (1980) and EPO App. Publ. No. 36,776) and the tac promoter (H. De Boer et al., *Proc. Natl. Acad. Sci. USA* **80**, 21 (1983). The promoter and Shine-Dalgarno sequence (for prokaryotic host expression) are operably linked to the DNA of the present invention, i.e., they are positioned so as to promote transcription of the messenger RNA from the DNA.

Expression vectors should contain a promoter which is recognized by the host organism. This generally means a promoter obtained from the intended host. While these are commonly used, other microbial promoters are suitable. Details concerning nucleotide sequences of many have been published, enabling a skilled worker to operably ligate them to DNA encoding the protein in plasmid or viral vectors (Siebenlist et al., *Cell* **20**, 269 (1980). The promoter and Shine-Dalgarno sequence (for prokaryotic host expression) are operably linked to the DNA encoding the desired protein, i.e., they are positioned so as to promote transcription of the protein messenger RNA from the DNA.

Eukaryotic microbes such as yeast cultures may be transformed with suitable protein-encoding vectors. See e.g., U.S. Patent No. 4,745,057. *Saccharomyces cerevisiae* is the most commonly used among lower eukaryotic host microorganisms, although a number of other strains are commonly available. Yeast vectors may contain an origin of replication from the 2 micron yeast plasmid or anautonomously replicating sequence (ARS), a promoter, DNA encoding the desired protein, sequences for polyadenylation and transcription termination, and a selection gene. An exemplary plasmid is YRp7, (Stinchcomb et al., *Nature* **282**, 39 (1979); Kingsman et al., *Gene* **7**, 141 (1979); Tschemper et al., *Gene* **10**, 157 (1980). This plasmid contains the trp1 gene, which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example

ATCC No. 44076 or PEP4-1 (Jones, *Genetics* **85**, 12 (1977). The presence of the trp1 lesion in the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoting sequences in yeast vectors include the promoters for metallothionein, 3-phospho-glycerate kinase (Hitzeman et al., *J. Biol. Chem.* **255**, 2073 (1980) or other glycolytic enzymes (Hess et al., *J. Adv. Enzyme Reg.* **7**, 149 (1968); and Holland et al., *Biochemistry* **17**, 4900 (1978), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., EPO Publn. No. 73,657.

Cultures of cells derived from multicellular organisms are a desirable host for recombinant protein synthesis. In principal, any higher eukaryotic cell culture is workable, whether from vertebrate or invertebrate culture, including insect cells. Propagation of such cells in cell culture has become a routine procedure. See Tissue Culture, Academic Press, Kruse and Patterson, editors (1973). Examples of useful host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, and WI138, BHK, COS-7, CV, and MDCK cell lines. Expression vectors for such cells ordinarily include (if necessary) an origin of replication, a promoter located upstream from the gene to be expressed, along with a ribosome binding site, RNA splice site (if intron-containing genomic DNA is used), a polyadenylation site, and a transcriptional termination sequence.

The transcriptional and translational control sequences in expression vectors to be used in transforming vertebrate cells are often provided by viral sources. For example, commonly used promoters are derived from polyoma, Adenovirus 2, and Simian Virus 40 (SV40). See, e.g., U.S. Patent No. 4,599,308. The early and late promoters are useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication. See Fiers et al., *Nature* **273**, 113 (1978). Further, the protein promoter, control and/or signal sequences, may also be used, provided such control sequences are compatible with the host cell chosen.

An origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral source (e.g. Polyoma, Adenovirus, VSV, or BPV), or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient.

Host cells such as insect cells (e.g., cultured *Spodoptera frugiperda* cells) and expression vectors such as the baculovirus expression vector (e.g., vectors derived from *Autographa californica* MNPV, Trichoplusia ni MNPV, *Rachiplusia ou* MNPV, or *Galleria ou* MNPV) may be employed to make proteins useful in carrying out the present invention, as described in U.S. Patents Nos. 4,745,051 and 4,879,236 to Smith et al. In general, a baculovirus expression vector comprises a baculovirus genome containing the gene to be expressed inserted into the polyhedrin gene at a position ranging from the polyhedrin transcriptional start signal to the ATG start site and under the transcriptional control of a baculovirus polyhedrin promoter.

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding DsrA may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing DsrA in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* **81**:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Rather than using vectors which contain viral origins of replication, one can transform mammalian cells by the method of cotransformation with a selectable marker and the chimeric protein DNA. An example of a suitable selectable marker is dihydrofolate reductase (DHFR) or thymidine kinase. See U.S. Pat. No. 4,399,216. Such markers are proteins, generally enzymes, that enable the identification of transformant cells, i.e., cells which are competent to take up exogenous DNA. Generally, identification is by survival or transformants in culture medium that is toxic, or from which the cells cannot obtain critical nutrition without having taken up the marker protein.

In general, those skilled in the art will appreciate that minor deletions or substitutions may be made to the amino acid sequences of peptides of the present invention without unduly adversely affecting the activity thereof. Thus, peptides containing such deletions or substitutions are a further aspect of the present invention. In peptides containing substitutions or replacements of amino acids, one or more amino acids of a peptide sequence may be replaced by one or more other amino acids wherein such replacement does not affect the function of that sequence. Such changes can be guided by known similarities between amino acids in physical features such as charge density, hydrophobicity/hydrophilicity, size and configuration, so that amino acids are substituted with other amino acids having essentially the same functional properties. For example: Ala may be replaced with Val or Ser; Val may be replaced with Ala, Leu, Met, or Ile, preferably Ala or Leu; Leu may be replaced with Ala, Val or Ile, preferably Val or Ile; Gly may be replaced with Pro or Cys, preferably Pro; Pro may be replaced with Gly, Cys, Ser, or Met, preferably Gly, Cys, or Ser; Cys may be replaced with Gly, Pro, Ser, or Met, preferably Pro or Met; Met may be replaced with Pro or Cys, preferably Cys; His may be replaced with Phe or Gln, preferably Phe; Phe may be replaced with His, Tyr, or Trp, preferably His or Tyr; Tyr may be replaced with His, Phe or Trp, preferably Phe or Trp; Trp may be replaced with Phe or Tyr, preferably Tyr; Asn may be replaced with Gln or Ser, preferably Gln; KGln may be replaced with His, Lys, Glu, Asn, or Ser, preferably Asn or Ser; Ser may be replaced with Gln, Thr, Pro, Cys or Ala; Thr may be replaced with Gln or Ser, preferably Ser; Lys may be replaced with Gln or Arg; Arg may be replaced with Lys, Asp or Glu, preferably Lys or Asp; Asp may be replaced with Lys, Arg, or Glu, preferably Arg or Glu; and Glu may be replaced with Arg or Asp, preferably Asp. Once made, changes can be routinely screened to determine their effects on function with enzymes.

As noted above, the present invention provides isolated and purified DsrA proteins, such as mammalian (or more preferably human) DsrA. Such proteins can be purified from host cells which express the same, in accordance with known techniques, or even manufactured synthetically.

Nucleic acids of the present invention, constructs containing the same and host cells that express the encoded proteins are useful for making proteins of the present invention. Specific initiation signals may also be used to achieve more

efficient translation of polynucleotide sequences encoding DsrA. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding DsrA, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. *Results Probl. Cell Differ.* **20**,125-162(1994)). In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; Bethesda, Md.) and may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express DsrA may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al., *Cell* **11**, 223-32 (1977)) and adenine phosphoribosyltransferase (Lowy, I. et al., *Cell* **22**, 817-23 (1980)) genes which can be employed in tk- or aprt- cells, respectively. Also, antimetabolite or antibiotic resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al., *Proc. Natl. Acad. Sci.* **77**, 3567-70 (1980)); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. et al., *J. Mol. Biol.* **150**,1-14 (1981)) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, supra). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest (i.e., *dsrA*) is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding DsrA is inserted within a marker gene sequence, transformed cells containing sequences encoding DsrA can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding DsrA under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain the nucleic acid sequence encoding DsrA and express DsrA may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

As explained further herein, proteins of the present invention are useful as immunogens for making antibodies as described herein, and these antibodies and proteins provide a "specific binding pair." Such specific binding pairs are useful as components of a variety of immunoassays and purification techniques, as is known in the art. The proteins of the present invention are of known amino acid sequence as disclosed herein, and hence are useful as molecular weight markers in determining the molecular weights of proteins of unknown structure.

The presence of polynucleotide sequences encoding DsrA can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding DsrA. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences encoding DsrA to detect transformants containing DNA or RNA encoding DsrA.

A variety of protocols for detecting and measuring the expression of DsrA, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on DsrA is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; *Serological Methods, a Laboratory Manual*, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med.* **158**:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding DsrA include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding DsrA, or any fragments thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison

Wis.); and U.S. Biochemical Corp., Cleveland, Ohio)). Suitable reporter molecules or labels, which may be used for ease of detection, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding DsrA may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode DsrA may be designed to contain signal sequences which direct secretion of DsrA through a prokaryotic or eukaryotic cell membrane. Other constructions may be used to join sequences encoding DsrA to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and DsrA may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing DsrA and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al., *Prot. Exp. Purif.* **3**, 263-281 (1992)) while the enterokinase cleavage site provides a means for purifying DsrA from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al., *DNA Cell Biol.* **12**, 441-453 (1993)).

In addition to recombinant production, fragments of DsrA may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J., *J. Am. Chem. Soc.* **85**, 2149-2154 (1963)). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for

example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of DsrA may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

Antibodies that specifically bind DsrA (*i.e.*, antibodies which bind to a single antigenic site or epitope on the proteins) are useful for a variety of diagnostic and therapeutic purposes. Antibodies to DsrA may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with DsrA or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

Monoclonal antibodies to DsrA may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) *Nature* **256**:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* **81**:31-42; Cote, R. J. *et al.* (1983) *Proc. Natl. Acad. Sci.* **80**:2026-2030; Cole, S. P. et al. (1984) *Mol. Cell Biol.* **62**:109-120). Briefly, the procedure is as follows: an animal is immunized with DsrA or immunogenic fragments or conjugates thereof. For example, haptenic oligopeptides of DsrA can be conjugated to a carrier protein to be used as an immunogen. Lymphoid cells (e.g. splenic lymphocytes) are then obtained from the immunized animal and fused with immortalizing cells (e.g. myeloma or heteromyeloma) to produce hybrid cells. The hybrid cells are screened to identify those which produce the desired antibody.

Human hybridomas which secrete human antibody can be produced by the Kohler and Milstein technique. Although human antibodies are especially preferred for treatment of human, in general, the generation of stable human-human hybridomas for long-term production of human monoclonal antibody can be difficult. Hybridoma production in rodents, especially mouse, is a very well established procedure and thus, stable murine hybridomas provide an unlimited source of antibody of select characteristics. As an alternative to human antibodies, the mouse antibodies can be converted to chimeric murine/human antibodies by genetic engineering techniques. *See* V. T. Oi *et al.*, *Bio Techniques* **4(4)**:214-221 (1986); L. K. Sun *et al.*, *Hybridoma* **5** (1986).

The monoclonal antibodies specific for DsrA epitopes can be used to produce anti-idiotypic (paratope-specific) antibodies. See e.g., McNamara et al., Dec. 14, 1984, Science, page 1325; Kennedy, R. C. et al., (1986) Science 232:220. These antibodies resemble the DsrA epitope and thus can be used as an antigen to stimulate an immune response against *H. ducreyi*.

In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S. L. et al. (1984) *Proc. Natl. Acad. Sci.* **81,** 6851-6855; Neuberger, M. S. et al. (1984) *Nature* **312**:604-608; Takeda, S. et al. (1985) *Nature* **314**:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce DsrA-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton D. R. (1991) *Proc. Natl. Acad. Sci.* **88,**11120-3).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al., *Proc. Natl. Acad. Sci.* **86**, 3833-3837 (1989)); Winter, G. et al., (1991) *Nature* **349**, 293-299 (1991)).

Antibody fragments which contain specific binding sites for DsrA may also be generated. For example, such fragments include, but are not limited to, the

$F(ab')_2$ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W. D. et al. (1989) Science 254:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between DsrA and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering DsrA epitopes is preferred, but a competitive binding assay may also be employed (Maddox, supra).

Antibodies may be conjugated to a solid support suitable for a diagnostic assay (e.g., beads, plates, slides or wells formed from materials such as latex or polystyrene) in accordance with known techniques, such as precipitation. Antibodies may likewise be conjugated to detectable groups such as radiolabels (e.g., ^{35}S, ^{125}I, ^{131}I), enzyme labels (e.g., horseradish peroxidase, alkaline phosphatase), and fluorescent labels (e.g., fluorescein) in accordance with known techniques.

The proteins and peptides of this invention may be used as antigens in immunoassays for the detection of *H. ducreyi* in various tissues and body fluids e.g., blood, spinal fluid, sputum, etc. A variety of immunoassay systems may be used. These include: radio-immunoassays, ELISA assays, "sandwich" assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, fluorescent immunoassays, protein A immunoassays and immunoelectrophoresis assays.

In addition, nucleic acids having the nucleotide sequences of the gene encoding DsrA or any nucleotide sequences which hybridize therewith can be used as probes in nucleic acid hybridization assays for the detection of *H. ducreyi* in various tissues or body fluids of patients. The probes may be used in any nucleic any type of hybridization assay including: Southern blots (Southern, 1975, J. Mol. Biol. 98:508); Northern blots (Thomas et al., 1980, Proc. Nat'l Acad. Sci.

U.S.A. 77:5201-05); colony blots (Grunstein et al., 1975, Proc. Nat'l Acad. Sci. U.S.A. 72:3961-65), etc. Stringency of hybridization can be varied depending on the requirements of the assay. Assays for detecting the polynucleotides encoding DsrA in a cell, or the extent of amplification thereof, typically involve, first, contacting the cells or extracts of the cells containing nucleic acids therefrom with an oligonucleotide that specifically binds to DsrA polynucleotide as given herein (typically under conditions that permit access of the oligonucleotide to intracellular material), and then detecting the presence or absence of binding of the oligonucleotide thereto. Again, any suitable assay format may be employed (see, e.g., U.S. Patent No. 4,358,535 to Falkow et al.; U.S. Patent No. 4,302,204 to Wahl et al.; 4,994,373 to Stavrianopoulos et al; 4,486,539 to Ranki et al.; 4,563,419 to Ranki et al.; and 4,868,104 to Kurn et al.) (the disclosures of which applicant specifically intends be incorporated herein by reference).

Kits for determining if a sample contains proteins of the present invention will include at least one reagent specific for detecting the presence or absence of the protein. Diagnostic kits for carrying out antibody assays may be produced in a number of ways. In one embodiment, the diagnostic kit comprises (a) an antibody which binds proteins of the present invention conjugated to a solid support and (b) a second antibody which binds proteins of the present invention conjugated to a detectable group. The reagents may also include ancillary agents such as buffering agents and protein stabilizing agents, e.g., polysaccharides and the like. The diagnostic kit may further include, where necessary, other members of the signal-producing system of which system the detectable group is a member (e.g., enzyme substrates), agents for reducing background interference in a test, control reagents, apparatus for conducting a test, and the like. A second embodiment of a test kit comprises (a) an antibody as above, and (b) a specific binding partner for the antibody conjugated to a detectable group. Ancillary agents as described above may likewise be included. The test kit may be packaged in any suitable manner, typically with all elements in a single container along with a sheet of printed instructions for carrying out the test.

Antisense oligonucleotides and nucleic acids that express the same may be made in accordance with conventional techniques. See, e.g., U.S. Patent No. 5,023,243 to Tullis; U.S. Patent No. 5,149,797 to Pederson et al. The length of the

antisense oligonucleotide (i.e., the number of nucleotides therein) is not critical so long as it binds selectively to the intended location, and can be determined in accordance with routine procedures. In general, the antisense oligonucleotide will be from 8, 10 or 12 nucleotides in length up to 20, 30, or 50 nucleotides in length. Such antisense oligonucleotides may be oligonucleotides wherein at least one, or all, or the internucleotide bridging phosphate residues are modified phosphates, such as methyl phosphonates, methyl phosphonothioates, phosphoromorpholidates, phosphoropiperazidates and phosphoramidates. For example, every other one of the internucleotide bridging phosphate residues may be modified as described. In another non-limiting example, such antisense oligonucleotides are oligonucleotides wherein at least one, or all, of the nucleotides contain a 2' loweralkyl moiety (e.g., C_1-C_4, linear or branched, saturated or unsaturated alkyl, such as methyl, ethyl, ethenyl, propyl, 1-propenyl, 2-propenyl, and isopropyl). For example, every other one of the nucleotides may be modified as described. See also P. Furdon et al., *Nucleic Acids Res.* **17**, 9193-9204 (1989); S. Agrawal et al., *Proc. Natl. Acad. Sci. USA* **87**, 1401-1405 (1990); C. Baker et al., *Nucleic Acids Res.* **18**, 3537-3543 (1990); B. Sproat et al., *Nucleic Acids Res.* **17**, 3373-3386 (1989); R. Walder and J. Walder, *Proc. Natl. Acad. Sci. USA* **85**, 5011-5015 (1988).

In another embodiment of the invention, DsrA, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between DsrA and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to DsrA, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with DsrA, or fragments thereof, and washed. Bound DsrA is then detected by methods well known in the art. Purified DsrA can also be coated directly onto plates for use in the aforementioned drug screening

techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding DsrA specifically compete with a test compound for binding DsrA. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with DsrA.

The proteins, peptides, polynucleotides and vectors comprising the polynucleotides of the present invention may be used as immunogens in vaccines against *H. ducreyi*, which vaccines are an aspect of the present invention. When used as an immunogen, it is not necessary to use the entire DsrA protein, although the entire DsrA protein may be used. Polypeptides, fragments, and/or antigenic determinants of DsrA may also be used as immunogens in the practice of the invention. The vaccines are used to prevent or reduce susceptibility to *H. ducreyi* infection.

The vaccines comprise an immunologically effective amount of the immunogen in a pharmaceutically acceptable carrier. The combined immunogen and carrier may be an aqueous solution, emulsion, or suspension. An immunologically effective amount is determinable by means known in the art without undue experimentation, given the teachings contained herein. Pharmaceutically acceptable carriers are known to those skilled in the art and include stabilizers, diluents, and buffers. Suitable stabilizers include carbohydrates, such as sorbitol, lactose, mannitol, starch, sucrose, dextran, and glucose and proteins, such as albumin or casein. Suitable diluents include saline, Hanks Balanced Salts, and Ringers solution. Suitable buffers include an alkali metal phosphate, an alkali metal carbonate, or an alkaline earth metal carbonate.

The immunogens of the invention are immunogenic without adjuvant, however adjuvants may increase immunoprotective antibody titers or cell mediated immunity response. Such adjuvants could include, but are not limited to, Freund's complete adjuvant, Freund's incomplete adjuvant, aluminum hydroxide, aluminum phosphate, aluminum oxide or a composition that consists of a mineral oil, such as Marcol 52, or a vegetable oil and one or more emulsifying agents, dimethyldioctadecyl-ammonium bromide, Adjuvax (Alpha-Beta Technology),

Inject Alum (Pierce), Monophosphoryl Lipid A (Ribi Immunochem Research), MPL+ TDM (Ribi Immunochem Research), Titermax (CytRx), toxins, toxoids, glycoproteins, lipids, glycolipids, bacterial cell walls, subunits (bacterial or viral), carbohydrate moieties (mono-, di-, tri- tetra-, oligo- and polysaccharide) various liposome formulations or saponins. Other adjuvants that may be included in vaccine compositions of the present invention include, but are not limited to: surface active substances (e.g., hexadecylamine, octadecylamine, octadecyl amino acid esters, lysolecithin, dimethyl-dioctadecylammonium bromide), methoxyhexadecylglycerol, pluronic polyols; polyamines (e.g., pyran, dextransulfate, poly IC, carbopol); and peptides (e.g., muramyl dipeptide, dimethylglycine, tuftsin). The immunogen may also be incorporated into liposomes, or conjugated to polysaccharides and/or other polymers for use in a vaccine formulation. Combinations of various adjuvants may be used with the conjugate to prepare the immunogen formulation. Exact formulation of the vaccine compositions will depend on the particular conjugate, the species to be immunized and the route of administration.

The vaccines of the invention are prepared by techniques known to those skilled in the art, given the teachings contained herein. Generally, the immunogens are mixed with the carrier to form a solution, suspension, or emulsion. One or more of the additives discussed above may be in the carrier or may be added subsequently. The vaccine preparations may be dessicated, for example, by freeze drying for storage purposes. If so, they may be subsequently reconstituted into liquid vaccines by the addition of an appropriate liquid carrier.

Any suitable vaccine and method of vaccination (i.e., immunization) known in the art may be employed in carrying out the present invention, as long as an active immune response against the antigen is elicited. When administered according to the present invention, the vaccine induces an active and protective immune response against unmodified cancer cells. Exemplary vaccination methods include, but are not limited to, "naked DNA" vaccines, viral and bacterial vector vaccines, liposome associated antigen vaccines, and peptide vaccines. Vaccines may be live vaccines, attenuated vaccines, killed vaccines, or subunit vaccines. Methods of vaccinating animals and humans against immunogens are well-known

in the art. *See, e.g.*, S. Crowe *et al.*, *Infections of the Immune System, in Basic and Clinical Immunology*, 697-715 (D. P. Stites & A. I. Terr, eds., 7th ed. 1991).

The vaccines of the present invention are administered to humans or other mammals, including bovine, ovine, caprine, equine, leporine, porcine, canine, feline and avian species, with humans being particularly preferred. The vaccines may administered to human children, including children younger than 18 months of age. Preferably, the vaccines of the present invention are administered to those subjects that are at particular risk of developing *H. ducreyi* infection (*i.e.*, subjects living in geographic locations where *H. ducreyi* is common).

The vaccines may be administered in one or more doses. The vaccines may be administered by known routes of administration for this type of vaccine, including parenteral administration, such as subcutaneous, intramuscular, or intravenous administration. Oral administration may also be used, including oral dosage forms which are enteric coated.

The schedule of administration of the vaccine may vary depending on the strain of *H. ducreyi* being used, the age and/or condition of the subject to be immunized, the particular formulation of the vaccine, and other factors known to those in the art. Subjects may receive a single dose, or may receive a booster dose or doses. Annual boosters may be used for continued protection.

The immunogens of this invention can be formulated as univalent and multivalent vaccines. The immunogens (i.e., the protein DsrA) can be mixed, conjugated or fused with other antigens, including B or T cell epitopes of other antigens. In addition to its utility as a primary immunogen, DsrA can be used as a carrier protein to confer or enhance immunogenicity of other antigens.

When a haptenic peptide of DsrA is used, (*i.e.*, a peptide which reacts with cognate antibodies, but cannot itself elicit an immune response), it can be conjugated to an immunogenic carrier molecule. For example, an oligopeptide containing one or more epitopes of DsrA may be haptenic. Conjugation to an immunogenic carrier can render the oligopeptide immunogenic. Preferred carrier proteins for the haptenic peptides of DsrA are tetanus toxin or toxoid, diphtheria toxin or toxoid and any mutant forms of these proteins such as CRM197. Others include exotoxin A of *Pseudomonas*, heat labile toxin of *E. coli* and rotaviral particles (including rotavirus and VP6 particles). Alternatively, a fragment or

epitope of the carrier protein or other immunogenic protein can be used. For example, the hapten can be coupled to a T cell epitope of a bacterial toxin.

The peptides or proteins of this invention can be administered as multivalent subunit vaccines in combination with other antigens of *H. ducreyi*. For example, they may be administered in conjunction with oligo- or polysaccharide capsular components of *H. ducreyi* such as polyribosylribitolphosphate (PRP).

Peptides and proteins having epitopes of DsrA evoke bactericidal antibodies which may act synergistically in killing *H. ducreyi* with antibodies against other outer membrane proteins of *H. ducreyi*. Thus, in an embodiment of the invention, DsrA (or a peptide or protein having a common epitope) is administered in conjunction with other outer membrane proteins of *H. ducreyi* (or peptides or proteins having epitopes thereof) to achieve a synergistic bactericidal activity. For combined administration with epitopes of other outer membrane proteins, the DsrA peptide can be administered separately, as a mixture or as a conjugate or genetic fusion peptide or protein. The conjugates can be formed by standard techniques for coupling proteinaceous materials. Fusions can be expressed from fused gene constructs prepared by recombinant DNA techniques as described.

The immunogens of this invention can be administered as live vaccines. To this end, recombinant microorganisms are prepared that express the peptides or proteins. The vaccine recipient is inoculated with the recombinant microorganism which multiplies in the recipient, expresses the DsrA peptide or protein and evokes a immune response to *H. ducreyi*. Preferred live vaccine vectors are pox viruses such as vaccinia (Paoletti and Panicali, U.S. Pat. No. 4,603,112) and attenuated Salmonella strains (Stocker, U.S. Pat. No. 4,550,081).

Live vaccines are particularly advantageous because they lead to a prolonged stimulus which can confer substantially long-lasting immunity. When the immune response is protective against subsequent *H. ducreyi* infection, the live vaccine itself may be used in a preventative vaccine against *H. ducreyi*.

Multivalent live vaccines can be prepared from a single or a few recombinant microorganisms that express different epitopes of *H. ducreyi*. In addition, epitopes of other pathogenic microorganisms can be incorporated into the vaccine. For example, a vaccinia virus can be engineered to contain coding

sequences for other epitopes in addition to those of *H. ducreyi*. Such a recombinant virus itself can be used as the immunogen in a mulivalent vaccine. Alternatively, a mixture of vaccinia or other viruses, each expressing a different gene encoding for different epitopes of outer membrane proteins of H. influenza and/or epitopes of other disease causing organisms can be formulated as a multivalent vaccine.

An inactivated virus or bacterial vaccine may be prepared. Inactivated vaccines are "dead" in the sense that their infectivity has been destroyed, usually by chemical treatment (e.g., formaldehyde treatment). Ideally, the infectivity of the virus or bacteria is destroyed without affecting the proteins which carry the immunogenicity of the vector. In order to prepare inactivated vaccines, large quanitites of the recombinant vector expressing the desired epitopes are grown in culture to provide the necessary quantity of relevant antigens. A mixture of inactivated viruses or bacteria expressing different epitopes may be used for the formulation of "multivalent" vaccines. In certain instances, these "multivalent" inactivated vaccines may be preferable to live vaccine formulation because of potential difficulties arising from mutual interference of live viruses administered together. In either case, the inactivated virus or mixture of viruses should be formulated in a suitable adjuvant in order to enhance the immunological response to the antigens. Suitable adjuvants include: surface active substances, e.g., hexadecylamine, octadecyl amino acid esters, octadecylamine, lysolecithin, dimethyl-dioctadecylammonium bromide, N, N-dicoctadecyl-N′-N′bis (2-hydroxyethyl-propane diamine), methoxyhexadecylglycerol, and pluronic polyols; polyamines, e.g., pyran, dextransulfate, poly IC, carbopol; peptides, e.g., muramyl dipeptide, dimethylglycine, tuftsin; oil emulsions; and mineral gels, e.g., aluminum hydroxide, aluminum phopshate, etc.

One particularly preferred embodiment of the invention is an attenuated vaccine comprising an *H. ducreyi* strain that does not express DsrA. The *H. ducreyi* strains that do not express DsrA used in these vaccines may be naturally occurring strains, or may be recombinant and/or isogenic mutants of *H. ducreyi* strains that do express the protein. Of these attenuated vaccines, a vaccine comprising the *H. ducreyi* mutant strain FX517 described herein is most preferred.

The bactericidal antibodies induced by DsrA epitopes can be used to passively immunize an individual against *H. ducreyi*. Passive immunization

confers short-term protection for a recipient by the administration of the pre-formed antibody. Passive immunization can be used on an emergency basis for special risks, e.g., young children exposed to contact with subjects already afflicted with *H. ducreyi* infection (chancroid).

In view of the foregoing description, the invention also comprises a method for inducing an immune response to *H. ducreyi* in a mammal in order to protect the mammal against infection by invasive or non-invasive *H. ducreyi*. The method comprises administering an immunologically effective amount of the immunogens of the invention to the host and, preferably, administering the vaccines of the invention to the host.

The following Examples are provided to illustrate the present invention, and should not be construed as limiting thereof. Unless otherwise noted, all chemicals and reagents were from Sigma Chemicals (St. Louis, MO). Standard recombinant DNA methods were used as described in Sambrook *et al.* (*supra*) or following manufacturers instructions.

EXAMPLE 1

Materials and Methods: Bacterial Strains and Media

Bacterial strains used in the experiments described herein are shown below in **Table 1**. For routine growth, *H. ducreyi* was maintained on chocolate agar plates obtained from UNC Hosptial Clinical Microbiology Lab. This medium was prepared using Mueller Hinton base and contained no fetal calf serum. When 5% fetal calf serum was required for optimal growth (*H. ducreyi* strains CHIA and 1157), Gonococcal medium base (GCB) used for preparation and instructions were followed (Difco). Antibiotics were used at the following concentrations for *E. coli*: ampicillin, 100 μg/ml; chloramphenicol, 30 μg/ml; kanamycin, 30 μg/ml; μg/ml; streptomycin, 100 μg/ml. For *H. ducreyi*, antibiotics were chloramphenicol, 1 μg/ml or streptomycin, 100 μg/ml.

Table 1. Bacterial strains and plasmids

Strain/Plasmid	Relevant Genotype/Phenotype	Source/Reference/Isolated
E. coli K-12		
DH5αMCR	*recA, gyrB*	Bethesda Research Labs
H. ducreyi		
35000	wild type	Stanley Spinola Indiana Univ.
FX516	35000 Co-integrate beta galactosidase positive intermediate in FX517 construction, Cm^r	This work
FX517	35000 *dsrA*, Cm^r	This work
CIP542 (Canada)		William Albritton
CIP A77		Robert Munson
CIP 542 (CDC)		Stephen Morse Centers for Disease Control
H. ducreyi obtained from Pat Totten		**(10)**
CIP A75		Pasteur Institute
CHIA		VDRL
HD167		VDRL
V-1157		Seattle
V-1168		Seattle
M90-02		Bahamas
406		Mississippi
425		Mississippi
54		Mississippi
010-2		Dominican Republic
HD301		Thailand
HD350		Kenya

Table 1. Bacterial strains and plasmids continued

Strain/Plasmid	Relevant Genotype/Phenotype	Source/Reference/Isolated
Plasmids		
pCRII	PCR cloning vector Kan^r, Amp^r	Invitrogen
pUNCH 1248	*dsrA* PCR clone using primers 14 and 16 in pCRII vector	This work
pLS88	Shuttle plasmid Kan^r, Str^r, Sul^r	(9)
pUNCH 1254	*dsrA* subclone. *ECo*R1 fragment of pUNCH 1248 in . *Eco*R1 of pLS88	This work
pUNCH 1255	mutagenized *dsrA*; pUNCH 1254 mutagenized with CAT cassette from pNC40 Kan^r, Cm^r	This work
pRSM1791	Mutagenesis plasmid Beta gal^r, Amp^r	This work
pUNCH 1256	pUNCH 1255 (SmaI/HinCII/Klenow) into the NotI (Klenow) of pRSM1791	This work
pUNCH 1260	*dsrA* PCR clone using primers 14 and 16 in pLSKS	This work
pNC40	source of CAT cassette, Amp^r,Cm^r	(37)

EXAMPLE 2

Outer membrane isolation, analysis, SDS-PAGE and immunoblotting

Large scale cultures of *H. ducreyi* were performed in Fernbach flasks with 1 liter of GCB-I broth containing 5% fetal calf serum and 50 μg/ml heme (Elkins, C. *Identification and purification of a conserved heme-regulated hemoglobin-binding outer membrane protein from Haemophilus ducreyi. Infec Immun.* **63,** 1241-1245 (1995)). Cultures of *E. coli* were performed in LB broth or LB agar plates containing appropriate antibiotics. Outer membranes were harvested as previously described **Id**. Protein concentrations were determined using the BCA kit

from Pierce (Rockford, IL). SDS-PAGE and Western blotting were performed as previously described (11). The lipooligosaccharide (LOS) of *H. ducreyi* was prepared using the method of Hitchcock and Brown (Hitchcock, P.G., and Brown, T.M., *Morphological heterogeneity among Salmonella LPS chemotypes in silver-stained polyacrylamide gels. J. Bacteriol.* **154**, 269-277 (1983). LOS was analyzed by SDS-PAGE and silver staining (Tsai, C.M. and Frasch, C.E., *A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal. Biochem.* **155**, 115-119 (1982)) or Western blotting with Mab 3F11 (Apicella, M.A. et al., *Phenotypic variation in epitope expression of the Neisseria gonorrhoeae lipooligosaccharide. Infect Immun.* **55**:1755-1761 (1987).

EXAMPLE 3

N-terminal sequence amino acid (aa) determination

The N-terminal aa sequence of DsrA was determined from strain 35000. Outer membranes were subjected to preparative SDS-PAGE and Western transfer to PVDF. The blot was stained temporarily with Ponceau S protein stain to locate the DsrA protein, which in strain 35000 migrates just below the 30 kDa standard protein. Strips of the blot were probed with anti-OpaF (generously provided by Janice Babcock and Richard Rest of Hahnemann Medical College) of gonococcal strain FA1090 and Mab 5C9. Anti-OpaF, for unknown reasons, cross-reacts with DsrA and Mab 5C9 reacts with a previously described *H. ducreyi* lipoprotein (termed Hlp) of similar molecular weight (18). These antibodies were used in order to unequivocally identify the proper band to sequence. The corresponding 30kDa-OpaF reactive band from the remainder of the Ponceau S stained blot was sequenced. The sequence obtained from the 30 kDa band was QQPPKFAGVS SLYSYEYDYG KGKKTKSNEG (amino acid residues 22–51, **SEQ ID NO:2**). This sequence did not match the processed mature, N-terminal sequence of Opa or Hlp 28 kDa (Hlp would be expected not to sequence, since it is a lipoprotein). We concluded that these three proteins were distinct.

The antiserum to DsrA was produced as follows. Outer membranes from *H. ducreyi* strain 35000 were electophoresed on large preparative well 12% SDS-PAGE gels. The gel was briefly stained and the corresponding 30 kDa band excised and electroeluted using a Centrilutor (Amicon) following the manufactuers

instructions. Mice were immunized a total of 3 times with 25 μg of gel purified protein per immunization. Freunds complete adjuvant was used for the first immunization and incomplete for the remainder.

EXAMPLE 4

Vector-Anchored PCR

Two degenerate oligonucleotides deduced from the N-terminal amino acid sequence (#6 and #7, **Fig. 2**) specifically hybridized to a 1.1 kb *Eco*R1 genomic fragment (data not shown). Attempts to clone this fragment using size selected DNA using several plasmid vectors were unsuccessful. Therefore a series of three separate vector-anchored PCR strategies were utilized to clone the *dsrA* structural gene, upstream flanking DNA and downstream flanking DNA, respectively. The first vector-anchored PCR (Fig. 2, V-A PCR 1) used the ligation between the 1.1 kb *Eco*R1 size-selected DNA and vector pBluescript as template and used 5' primer #6 and vector primer KS as amplimers. An approximate 700 bp fragment was amplified and preliminary sequence obtained. The N-terminal sequence originally obtained from Edman degradation matched the deduced amino acid sequence of the PCR product, but was not homologous to known sequences in the data bases. In contrast, the C-terminus of the gene was homologous to UspA2 and YadA (see results below), this suggested the possibility of PCR generated artifact(s). To rule out PCR artifact additional PCR was performed. The primers used included 5' primers #6, 8 and 9 and 3' primers 11 and 12. The latter 4 primers were derived from the DNA sequence obtained from the original anchored PCR product above (**Fig. 2** and data not shown). Identically sized products from total *H. ducreyi* chromosomal DNA template (and the original anchored PCR product, the + control template were amplified) using 3' primers from the region with homology to C-terminal YadA (primers #11 and #12) (data not shown). Furthermore, Southern hybridization of *H. ducreyi* chromosomal DNA probed with oligos #6, #7, #8, #9, #11, #12 and the PCR product generated from #8 and #12 all specifically recognized the 1.1 kb band *ECo*R1 band (**Fig. 1** and data not shown). It was concluded that the N-terminal aa sequence obtained from the 30 kDa protein is found on the same ORF that has C-terminal homology to UspA2/YadA. These data established that the open reading frame (ORF) data were correct.

To obtain sequence upstream of the structural gene for *dsrA*, a second vector-anchored PCR was used (**Fig. 2**, V-A PCR 2). Again, the template was the ligation between the 1.1 kb *Eco*R1 size-selected DNA and vector pBluescript but the primers used were #12 and vector primer KS. A (1069) bp fragment which included the upstream *Eco*R1 site (**Fig 2.**) was amplified.

To obtain sequence downstream of the *dsrA* gene a third vector-anchored PCR was used (**Fig. 2**, V-A PCR 3). Southern hybridization identified an approximate 4 kb *Bgl* II fragment which hybridized with *dsrA* probes and there are no *Bgl*II sites in the 1.1 kb *Eco*R1 fragment. Fragments of 3-5 kb *Bgl* II restricted chromosomal DNA were isolated and ligated to *Bam*H1, shrimp alkaline phosphatase treated pMCL210 vector. The ligaton reaction was ethanol precipitated and amplified using primers 10 and vector primer T7 (promoter), yielding an approximately 2.5 kb PCR product. The products of all three vector-anchored PCR reactions were sequenced with appropriate primers to obtain preliminary sequence and these sequences confirmed one another (data not shown).

Commercially available PCR tubes (Ready to Go, Pharmacia) were utilized for PCR. Analytical PCR (25 ul final volume) utilized single tubes whereas preparative PCR combined the "beads" from 4 tubes into single tube (100 ul final volume). The $MgCl_2$ concentration in all PCR reactions was 4 mM. The first two vector anchored PCRs used 5 ul of ligation and 25 pm of each primer. The conditions for PCR for first two vector anchored PCRs were: hot start 5 min 94C, denature 94C; 1 minute annealing, 50C, 1 minute; extension 72, 1 minute. The conditions for the third PCR were identical except that the extension time was 3 min.

EXAMPLE 5

DNA sequencing and analysis.

DNA sequence analysis was performed at the University of North Carolina at Chapel Hill Automated Sequencing Facility utilizing Taq terminator chemistry. The final sequences presented for strain 35000 in **Fig. 2** and for the other *H. ducreyi* strains in **Fig. 9** was obtained from PCR products using primers #14 and 24 which flank the *dsrA* gene (**Fig.1**). Both strands of the were completely sequenced. The sequence data were assembled using the program AssemblyLIGN (IBI). The preliminary sequence for the *dsrA* structural gene from 35000 obtained

by vector-anchored PCR was in complete agreement with the final sequence presented (**Fig.3**). Amino acid alignments were done by Clustal in the program GeneJockeyII (Cambridge, UK) and PAM 250 setting. Bestfit (GCG Computer Group, Wisconsin) was used to generate similarity and identity scores using a gap weight of 8.

EXAMPLE 6

Plasmid Constructions

Plasmid pUNCH 1248 was constructed by PCR. A 900 bp fragment was amplified from *H. ducreyi* strain 35000 using primers 14 and 16 (**FIG. 2**) using the conditions described above for the first two vector anchored PCRs. The product was ligated to pCRII following manufacturers directions, tranformed into *E. coli* DH5□ and recombinants identified by restriction analysis. *E. coli* harboring pUNCH 1248 grew poorly, was propagated only on agar plates to reduce the possibilty of mutation/deletion, and gave rise to an occasional larger colony. Subclone 1254 was constructed by isolating the *Eco*R1 fragment of pUNCH 1248 and ligation into *Eco*R1 restricted pLS88. *dsrA* of pUNCH 1254 was mutagenized by insertion of a CAT (Chloramphenicol Acetyl Transferase) into the open reading frame to construct pUNCH 1255. To perform this, a CAT cassette (a *Bgl*II fragment from pNC40 was treated with Klenow to fill-in the ends) was ligated into the *Nde*I site of pUNCH 1254 (previously treated with Klenow to produce blunt ends). pUNCH 1256 was constructed by moving the insert from pUNCH 1255 (containing mutagenized *dsrA*) into plasmid pRSM1791 for subsequent mutagenesis. This was done by isolation of a *Sma*I to *Hin*CII fragment of pUNCH 1255, Klenow treatment and ligation into the *Not*I site of pRSM1791 previously treated with Klenow. Transformation of *E. coli* host was performed and selection using Amp and Cm yielded pUNCH 1256.

EXAMPLE 7

Construction and characterization of an *H. ducreyi dsrA* mutant

An isogenic mutant (FX517, Table 1) was constructed by allelic replacement of the wild-type locus of strain 35000 with the mutation in pUNCH 1256 using a previous system of mutagenesis described by Bozue et al (Bozue,

J.A. et al.; *Facile construction of mutations in Haemophilus ducreyi using lacZ as a counter-selectable marker. FEMS Microbiology Letters.* **164**, 269-73 (1998)). In this procedure, a mutagenized copy of the locus (containing a chloramphenicol (Cm or CAT) cassette) was subcloned into a plasmid able to express lacZ (pUNCH 1256). *H. ducreyi* were electroporated and Cm^r transformants selected (Elkins et al., *Characterization of the hgbA locus of Haemophilus ducreyi. Infect Immun.* **63**, 2194-2200 (1995); Hansen, E.J. et al., *Use of electroporation to construct isogenic mutants of Haemophilus ducreyi. J. Bacteriol.* **174**, 5442-9 (1992)). These transformants putatively contained the entire plasmid integrated due to a single crossover event (as exemplified by FX516, Table 1). Individual transformants were streaked onto Cm medium containing X-gal. Since the product of X-gal is highly toxic to *H. ducreyi* the co-integrates grow as tiny blue colonies. The loss of the X-gal sequences and neighboring wild type allele via a resolution of the co-integrate results in only the mutant allele being retained (exemplified by FX 517, Table 1). These *H. ducreyi* mutants grew as normal-sized white colonies on the medium containing Cm and X-gal similar to other *H. ducreyi* mutants containing CAT cassettes (Elkins, C. et al., *Characterization of the hgba locus of Haemophilus ducreyi. Infect Immun.* **63**, 2194-2200 (1995); Elkins, C. et al., *Role of the Haemophilus ducreyi Ton system in internalization of heme from hemoglobin. Infection & Immunity* **66**,151-60 (1998); Thomas, C. et al., *Cloning and characterization of tdhA, a locus encoding a TonB-dependent heme receptor from Haemophilus ducreyi. Infect Immun.* **66**, 1-9 (1998)) and data not shown.

Southern blot and PCR analysis was used to confirm that an allelic replacement occurred in the generation of *H. ducreyi* mutant FX517. Chromosomal DNA was isolated from strains 35000, FX516, and FX517, digested with *Hin*CII and subjected to electrophoresis and bidirectional transfer. The two blots were probed with either the PCR product of oligos 14 and 16 or the *Bgl* II CAT fragment from pUNCH 40. Digoxigenin-labeled, bound probe was detected with alkaline phosphatase labeled anti-digoxigenin antibody (Boehringer Mannheim) followed by detection with nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP). PCR confirmation of the mutant utilized primers 14 and 16 which flank the *Nde*I site (CAT cassette) used for gene disruption.

EXAMPLE 8

Complementation of FX517 and other *dsrA* mutants in *trans*

To rule out that the serum susceptibility of *dsrA* mutant FX517 was due to a mutation elsewhere on the chromosome or polar downstream effects, complementation in trans was performed. Briefly, we PCR amplified the *dsrA* and surrounding locus using primers 14 and 24 (**Fig. 2**), Klenow treated the PCR product, and restricted the PCR product with *Hin*DIII (which restricts just downstream of *dsrA*, **Fig.2**). After gel purification, the PCR product was ligated into *Sma*I/*Hin*DIII restricted pLSKS (Wood, G.E. et al., *Target and cell range of the Haemophilus ducreyi hemolysin and its involvement in invasion of human ipithelial cells. Infect and Immun.* **In Press.**) The ligation was ethanol precipitated and *H. ducreyi* strain FX517 electroporated. Streptomycin resistant colonies were screened for production of DsrA by Western blotting and confirmed by restriction analysis. One experimental transformant, pUNCH 1260*dsrA*, and one vector transformant were selected for further study. pUNCH 1260 and the vector pLSKS (negative control) were then electroporated into the three naturally occurring *dsrA* mutants (CIP A75, CIP A77, CIP 542 (Can), Table 1).

EXAMPLE 9

Serum susceptibility

The resistance of *H.ducreyi* to normal human serum was performed as previously described (Odumeru; Carbonetti) with the following modifications: An 18-24 hour culture of *H. ducreyi* from chocolate agar plates was scraped into GCB broth to an OD600 of 0.2. A 10^{-4} to 10^{-5} dilution was made (approximately 1000 CFU/ml, depending on the strain) and aliquots mixed with pooled fresh normal human serum (NHS) or heat inactivated NHS (56C, 30 min) to a final concentration of 25 or 50% NHS. After incubation for 45 minutes at 35C in 5% CO2, 100 ul aliquotes were plated onto chocolate agar plates and viable counts performed after 48 hours. Data are expressed as percent survival in the fresh NHS as compared to survival in heat-inactivated NHS (number of CFU survivors in fHNS/number of survivors in heated NHS X 100). Strains containing pUNCH 1260 or pLSKS were propagated and plated on chocolate agar containing streptomycin at 100 μg/ml.

EXAMPLE 10

Identification of a 30 kDa protein involved in serum resistance.

During the course of studies characterizing the *H. ducreyi* interaction with PMNs, a series of Western blots were performed using various antibodies to the Opa proteins from *gonococci*. It was found that a polyclonal antiserum to OpaF of gonococcal strain FA1090 reacted at a dilution of 1:5000 with a protein (DsrA) that varied between 28 and 35 kDa in a panel of strains (data not shown). One strain, CIPA75, did not react. CIPA75 was of interest because it had previously been shown to be avirulent in the chilled rabbit model of infection, to be serum susceptible, to exhibit reduced adherence to HEp-2 cells and to have a truncated LOS (Odumeru, J.A. et al, *Role of lipopolysaccharide and complement in susceptibility of Haemophilus ducreyi to human serum. Infectt Immun.* **50**,495-9 (1985); Rice, P.A., *Molecular basis for serum resistance in Neisseria gonorrhoeae. Clinical Microbiology Reviews.* **2**, S112-7 (1989). Specific antisera to DsrA was generated using DsrA purified by preparative SDS-PAGE and electroelution of outer membranes from *H. ducreyi* strain 35000. Western blots of several geographically diverse lab and clinical isolates were probed with anti-DsrA (**Fig. 1**). This was done to confirm that the previous cross-reactivity seen with the anti-OpaF serum was due to the presence of DsrA and to ascertain what percentage of strains expressed *dsrA*. The proteins recognized in the DsrA Western blot (**Fig. 1**) and the OpaF Western blot (data not shown) appeared to be identical. Most strains in **Fig. 1**, expressed an immunoreactive protein, except for the previously reported avirulent strains CIP A75, CIP A77 (25-27) and CIP542 (Can., obtained from Canada) (Alfa, M.J. et al., *Use of tissue culture and animal models to identify virulence-associated traits of Haemophilus ducreyi. Infection & Immunity* **63**:1754-61 (1995)). In contrast, virulent CIP 542 (CDC), obtained from the CDC and previously shown to cause a laboratory acquired infection (Trees, D.L. et al., *Laboratory-acquired infection with Haemophilus ducreyi type strain CIP 542. Med Microbiol.* 330-337 (1992)), expressed *dsrA*. Previous studies documented that virulent *H. ducreyi* strains are serum resistant. We performed serum susceptiblity studies of selected *H. ducreyi* strains which did and did not express *dsrA* and these results are summarized at the bottom of **Fig 1**. For the purposes of this study, we

arbitrarily termed a strain serum resistant if there were more than 10% survivors when exposed to 50% fNHS serum as compared to NHS. The specific percent survivors (+/- sd) for each of the strains tested in Figure 1 are: 35000, 79%; CIP A75; CIP A77; CIP 542 (Can); CIP 542 (CDC); CHIA; V-1157; M90-02; and 406. Thus, in these initial studies there was a correlation between strains tested which expressed detectable *dsrA* and serum resistance. This correlation between the lack expression of *dsrA* and serum susceptibility in the *dsrA* mutant strains, some of which also had LOS alterations, could merely be coincidental. Therefore additional molecular studies were performed culminating in the generation of an isogenic *dsrA* mutant for biological studies.

EXAMPLE 11

Molecular Studies

Through a series of experiments involving Western blotting, immunoprecipitation and finally N-terminal amino acid sequencing, it was determined that the DsrA protein was not the same as the previously described 28 kDa lipoprotein termed Hlp (17) (data not shown). The N-terminal amino acid sequence of the immunoreactive DsrA 30 kDa protein of strain 35000 was found to be: QQPPKFAGVS SLYSYEYDYG KGKKTKSNEG (amino acid residues 22–51, **SEQ ID NO:2**). No known homologies were initially detected when this peptide sequence was searched against GENBANK, including gonococcal Opa proteins.

Two degenerate oligonucleotides (#6 and #7) were synthesized based on the above N-terminal sequence and found to hybridize specifically to a 1.1 kB *EcoR*1 chromosomal band from *H. ducreyi* strain 35000 (data not shown). Attempts to clone this fragment were unsuccessful and three separate vector-anchored PCR reactions (V-A PCR) were used to amplify the relevant locus and surrounding regions (**Fig. 2**). Preliminary sequencing of the product of V-A PCR 1 **(Fig. 2)** identified an ORF that was homologous to the UspA2 protein of *Moraxella catarrhalis* and the YadA protein of *Yersinia spp.*, but only in the C-terminal region. Since both of these proteins are implicated in determining important virulence traits (including serum resistance), additional studies were undertaken.

EXAMPLE 12

DNA and deduced amino acid sequence of the *H. ducreyi dsrA* locus from strain 35000

The DNA sequence of the *dsrA* locus, including 100 bp of sequence upstream of the ATG start and 126 bp of sequence downstream of the TAA termination codon are presented in **Fig. 3**. The data presented were obtained from PCR products amplified using primers 14 and 24. Sequences similar to -35 (TGATAA) and -10 (TATATT) *E. coli* promoter consensus sequences were found beginning at nt 13 (TTGACA) and nt 35 (TAGAAT) respectively, and were separated by 16 nt. A putative ribosome-binding site (TAATGAGG) was found beginning 13 nt upstream of the *dsrA* start codon. Beginning at nt 913 and ending at nt 946 was an inverted hairpin loop containing 13 matched nucleotides, consistent with a transcription terminator. The gene immediately downstream of *dsrA* and in the opposite orientation was an ORF with homology to the hypothetical protein HI0107 of the genome sequence of *H. ducreyi*. The GC content of the 1 kb of DNA sequence presented was 34.5%, consistent with the AT-rich nature of *Haemophilus spp.* DNA.

The *dsrA* ORF predicted a protein of 28215 daltons, which when processed would yield a mature protein of 26375 daltons. This is in agreement with migration in SDS-PAGE for strain 35000 (**FIG. 1**). Comparison of the deduced amino acid sequence of DsrA with the N-terminal amino acid sequence revealed identity in 28 of 30 amino acids. The first two residues of the mature protein, QQ, were unusual in their charges; however, certain versions of mature YadA begin with two charged amino acids (see below). Just preceding the DsrA QQ residues was the unusual signal peptidase I cleave site of TMA. Consistent with the outer membrane localization, DsrA contained a carboxyl terminal motif ending with a phenylalanine which is found in the majority of integral outer membrane proteins (Struyve, M. et al., *Carboxyl-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J. Mol. Biol.* **218**, 141-148 (1991)). The mature DrsA protein was predicted to be very basic with a pI of 9.1 and which accounts for its poor transfer during Western blotting (data not shown).

Alignment of the DsrA protein with similar proteins is shown in **FIG. 4**. DsrA was most similar to UspA2 and YadA in a region of the C-terminus and was most divergent in the N-terminus. Using the Bestfit program, DsrA was 45% similar and 40% identical to UspA2; DsrA was 47% similar and 39% identical to YadA. It should be noted that both of these heterologous proteins are considerable larger than DsrA which may account for such differences in the N-terminal domains. The C-terminus of YadA is believed to be anchored in the outer membrane and the N-terminus encodes the functional regions of the YadA protein (Rogenkamp, A. et al., *Substitution of two histidine residues in YadA protein of Yersinia enterocolitica abrogates collagen binding, cell adherence and mouse virulence. Molecular Microbiology* **16**, 1207-19 (1995); Roggenkamp, A. et al., *Deletion of amino acids 29 to 81 in adhesion protein YadA of Yersinia enterocolitica serotype 0:8 results in selective abrogation of adherence to neutrophils. Infection & Immunity* **65**, 2506-14 (1996); Tamm, A., et al., *Hydrophobic domains affect the collagen-binding specificity and surface polymerization as well as the virulence potential of the YadA protein of Yersinia enterocolitica. Molecular Microbiology.* **10**, 995-1011 (1993)).

EXAMPLE 13

Construction and characterization of an *H. ducreyi dsrA* mutant.

An isogenic mutant (FX517, Table 1) was constructed by allelic replacement of the wild-type locus of strain 35000. Initial attempts to obtain a double crossover with a CAT cassette in the cloned gene were unsuccessful using pUNCH 1255 (data not shown). Therefore, we used a recently described method to obtain mutants (Bozue, J.A. et al., *Facile construction of mutations in Haemophilus ducreyi using lacZ as a counter-selectable marker. FEMS Microbiology Letters.* **164**:269-73 (1998)).Using this procedure, several chloramphenicol resistant cointegrates were obtained. After streaking each cointegrate onto X-gal chocolate plates, several mutants were obtained for each cointegrate and none of mutants expressed *dsrA* (data not shown). One mutant, FX517 was selected for further study. Outer membranes were made from the parent and mutant strain FX517 and subjected to SDS-PAGE and Coomassie staining or SDS-PAGE and Western blotting (Figures 5A and 5B, respectively). DsrA is an abundant outer membrane protein in strain

35000 but is absent in the mutant. No reactivity was obtained from FX514 using anti-DsrA antisera (**FIG. 5**, Panel B) or anti-OpaF (data not shown). Similar to UspA2 and YadA, DsrA had a propensity to form multimers, especially when solubilized at the lower temperature of 37C (**FIG. 5**, Panel A, and data not shown).

The structure of the mutagenized *dsrA* locus in FX517 was confirmed using Southern blotting and PCR. In Southern blots of chromosomal DNA from the parent and mutant strains the *Hin*CII band recognized by the *dsrA* probe increased in size approximately 1 kb in the mutant as compared to the parent band. Similarly, an identical blot hybridized with the CAT probe recognized only the larger *Hin*CII band of the mutant (data not shown). PCR of the 35000 and FX517 *dsrA* locus with primers flanking the CAT insertion indicated the locus was approximately 1 kb larger in the mutant (data not shown). These data are consistent with an allelic replacement event.

EXAMPLE 14

Serum resistance phenotype of the *dsrA* mutant

The serum susceptibility of the naturally occurring *dsrA* mutants and the role of the related YadA and UspA2 proteins in mediating serum resistance prompted us to test FX517 for serum sensitivity. Serum killing studies of parent strain 35000 and *dsrA* mutant FX517 were performed using 25 and 50% normal pooled serum (Fig. 6). FX517 was very susceptible to NHS and demonstrated zero or 2% survival in 50% and 25% NHS, respectively. In contrast, parent strain 35000 was relatively serum resistant, exhibiting 79% and 50% survival in 50% and 25% NHS (p values 0.002 and 0.004 for 50% and 25% NHS, respectively, using Students paired T test). Thus DsrA appeared to be required for expression of a serum resistant phenotype.

EXAMPLE 15

Complementation of *dsrA* mutants

It was possible that a cryptic mutation had occurred during the construction of FX517 which could account for its serum susceptibility phenotype. Furthermore, we wished to determine whether the serum susceptibility of the three naturally occurring *dsrA* mutants could be converted to serum resistance if they

expressed *dsrA*. Each *dsrA* mutant (isogenic mutant FX517 or naturally occurring mutants CIP A75, CIP A77, and CIP 542 (Can)) was electroporated with pUNCH 1260 (*dsrA*) or pLSKS (vector control) plasmids. These shuttle plasmids are able to replicate in *H. ducreyi*. Strains containing pUNCH 1260, but not pLSKS, expressed *dsrA* (**Fig. 7A**). Subjectively, it appeared that more DsrA was expressed from the strains complemented with the *dsrA* plasmid than from 35000 (n=4), perhaps due to gene dosage or growth on medium containing streptomycin. Expression of *dsrA* from plasmid pUNCH 1260 suggested that the tentatively identified promoter (**Fig. 3**), was driving expression of the cloned *dsrA* gene since very little additional upstream DNA was present and the insert was in the opposite direction of the lac promoter in pLSKS.

Bactericidal killing was performed on each of the complemented *dsrA* mutants (**Fig. 7B**). For strains FX517, CIP A75, CIP A77, and CIP 542 (Can), expression of *dsrA* from pUNCH 1260 conferred serum resistance. However, for strains harboring the plasmid vector lacking an insert serum resistance was not conferred.

EXAMPLE 16

Lipooligosaccharide expression by *H. ducreyi*

In some bacterial systems, mutants in LOS are more serum susceptible. Indeed, it was reported by Odumeru that the reason for the serum susceptibility of *H. ducreyi* strains CIP A75 and CIP A77 was due to LOS truncation. It was possible that the lack of *dsrA* expression in *dsrA* mutants (FX517, CIP A75, CIP A77 and CIP 542 (CAN)) resulted in the truncation of LOS directly or indirectly. Alternatively repair of *dsrA* expression in LOS/*dsrA* apparent double mutants (CIP A75 and CIP A77) might affect LOS expression and subsequent serum susceptiblity. To address these possibilities, LOS was analyzed by SDS-PAGE and silver staining (**Fig 8**) and Western blotting (data not shown). We compared 35000 and FX517 LOS (without plasmids) in several silver stained gels and the migration patterns were always indistinguishable. Furthermore, Western blotting of 35000 and FX517 LOS with anti-LOS Mab 3F11 was similar.

Silver stained LOS gels of the complemented *dsrA* mutants were indistinguishable between each strain pair containing either pUNCH 1260 (*dsrA*)

or pLSKS, respectively. There was a minor variation in a faster migrating LOS band for some of the strains (CIP542, no plasmid present) when grown on antibiotic free chocolate (Mueller Hinton base) as compared to the same strain (CIP542, either plasmid present) grown on streptomycin chocolate (Gonococcal medium base). However, it should be noted that within each pair of matched strains (expressing or not expressing *dsrA*), there were no apparent major LOS differences. Thus, under the limited conditions examined here, the presence of DsrA and not LOS length was the dominant determinant of serum resistance.

EXAMPLE 17

Structural Analysis of *dsrA* in other *H. ducreyi* strains

Western blotting of a variety *H. ducreyi* strains (**Fig. 1**) suggested strongly that DsrA varied in molecular weight and/or amino acid sequence among the strains. Furthermore, we desired to understand whether mutations had occurred in the naturally occurring *dsrA* mutants or whether the possibility of phase variation could account for their inability to express *dsrA*. PCR was used to amplify a 1.2 kb fragment from 8 additional strains, including the *dsrA* mutants (**Fig. 2**, primers 14 and 24). The deduced amino acid sequence indicated that overall the DsrA protein was quite similar between strains (**Fig. 9**). Two regions with modest variability were observed and termed variable region 1 and 2 (VR1 and VR2). Variable region 1 included amino acids roughly 90-100 (depending on the strain) and a few substitutions and insertions were noted. Variable region 2 contained either 1, 2, or 3 identical copies of the heptamer repeat sequence NTHNINK (**SEQ ID NO:19**) and spanned amino acids 174–195 in the various strains. It is likely that the different number of repeat sequences was the predominant factor accounting for the variable migration seen in SDS-PAGE and Western blotting. Excepting for mutant strain CIP542(Can), which contained a stop codon (see below), the sequences for all other 8 DsrA proteins were identical after VR2. Thus, DsrA is highly conserved in sequence, despite its variable mobility in gels.

EXAMPLE 18

Affinity purification of DsrA using vitronectin (Vn)

H. ducreyi were surface iodinated using Iodogen-coated tubes as directed by the manufacturer. Briefly, to a tube coated with 50 µg of iodogen was added 0.5 mCi of NaI (Amersham IMS30) and 0.5 ml of 1 X 109.H. ducreyi in Phosphate buffered saline (PBS). After incubation for 2 minutes, the labeled bacteria were centrifuged and washed in medium to remove unincorporated iodine. The procedure labels primarily tyrosine residues on surfaced exposed proteins (outer membrane proteins). Each indicated biotinylated Vn was mixed with an aliquot of strain 35000 and strain FX517 whole surface-iodinated H. ducreyi. After Vn binding to *H. ducreyi* strains (15 min), unbound Vn was removed by centrifugation and washing. Bacteria and Vn were solubilized in the detergent ZW3,14 and insoluble material removed by centrifugation for 5 min at 15,000 X g. The detergent soluble proteins were mixed with strepavidin-agarose solid phase. After incubation (2 hours), extensive washing the strepavidin-agarose with ZW3,14 in PBS, the samples were boiled in Laemmli sample buffer, and subjected to SDS-PAGE and autoradiography. The results are shown in **FIG. 12**, showing that affinity purification of DsrA from whole cells is possible using biotinylated vitronectins (Vn).

EXAMPLE 19

Method of attachment of H. ducreyi to Human Cells.

Efficient attachment of *H. ducreyi* to a keratinocyte cell line requires DsrA expression. *H. ducreyi* were added to the HaCaT cells at a MOI of 10:1 and incubated for 4 hours. After removal of unbound bacteria by extensive washing, CFUs were determined by plating the disrupted monolayer. Results are shown in **FIG. 11,** illustrating showing that efficient attachment of *H. ducreyi* to a keratinocyte cell line requires DsrA expression. Data are from 4 experiments.

The foregoing is illustrative of the present invention and is not to be construed as limiting thereof. The invention is defined by the following claims, with equivalents of the claims to be included therein.

SEQUENCE LISTING

SEQ ID NO:1 DNA SEQUENCE OF 35000 (set forth in FIG. 3)

ATAAATACGTCATTGACATTTTT TTAATGTAAG GTAGAATAAG AAAGTAAATT
CTATATTTAC AATCAAGATT GACAATTATT TACTTAATGA GGTGATTATG
AAAATTAAAT GTTTAGTTGC CGTAGTGGGA TTAGCTTGTT CTACTATTAC
AACAATGGCT CAGCAGCCGC CAAAGTTTGC TGGAGTATCT TCTTTGTATA
GCTATGAGTATGACTATGGT AAGGGTAAAT GGACTTGGTC TAATGAAGGC
GGTTTCGATA TTAAAGTGCC AGGGATTAAA ATGAAGCCAA AAGAATGGAT
TTCTAAACAG GCTACTTATC TTGAATTACA GCATTATATG CCTTATACTC
CTGTTCTCGT GACATATGCT CCTGGCGTTT CTCCTAGCCC TATACTGTTA
TATCCGATGT CTGATCCTGA TCAACTTGGA ATAAATCGGC AGCAGCTGAA
ATTGAATTTG TATAGTTATT TTAACGATTT AAGACACGAT TTTAAATTAA
AAGTTCTTGA TGCACGTATT TCCAAAAATA AACAAAATAT TGATACTATA
AGTAAATATT TACTAGAACT GGGTACTTAT TTAGATGATT CTTATCGTAT
GATGGAACAA AATACACATA ATATCAATAA GTTGTCTAAA GAATTGCAAA
CTGGTTTAGC CAACCAATCA GCATTGTCTA TGTTAGTGCA ACCAAATGGT
GTAGGCAAAA CGAGCGTTTC TGCTGCGGTA GGAGGTTATA GAGATAAAAC
TGCATTAGCC ATTGGTGTCG GCTCACGCAT TACTGATCGC TTTACCGCTA
AAGCGGGTGT AGCGTTCAAT ACCTACAATG GCGGCATGTC TTATGGTGCT
TCTGTTGGTT ATGAATTCTA ATCATTACGT TTAATCACTA ATCGTTTTGG
TTATAATAAA AAGGCTAAAT GTTTCTCCTC ACATTTAGCC TTTCTTATTT
ATCTTTGTTA TAGCTTTTGC TGTTATAAAA CCGTTTTTTA GCCACTTTTA
TTAATTAAGC TTTTAAGCCT ATTCAATCAG TTCTACTTTC ACTTTTTTCA
CCATATTATC CGCCACTTCT AAAACGGTAA TATTAAGTTG GTTTAGCCTA
AATTGGGTAC CTTCTATCGG AATTTTTTCT AAATGTTCTA AAATTAAGCC
GTTAAAGGTG CGGAC

SEQ ID NO:2 PROTEIN SEQUENCE OF 35000 (set forth in FIG. 3)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLYSYEYDY GKGKWTWSNE
GGFDIKVPGI KMKPKEWISK QATYLELQHY MPYTPVLVTY APGVSPSPIL
LYPMSDPDQL GINRQQLKLN LYSYFNDLRH DFKLKVLDAR ISKNKQNIDT
ISKYLLELGT YLDDSYRMME QNTHNINKLS KELQTGLANQ SALSMLVQPN
GVGKTSVSAA VGGYRDKTAL AIGVGSRITD RFTAKAGVAF NTYNGGMSYG
ASVGYEF

SEQ ID NO:3 DNA SEQUENCE OF CIPA75

ATTTTATAATTTACAATACATTTTATATTTTTATATTATATAAATACGTCATTGACATTT
TTTTAAGGTAGAATAAGAAAGTAAATTCTATATTTACAATCAAGATTGACAATTATTTA
CTTAATGAGGTGATTATGAAAATTAAATGTTTAGTTGCCGTAGTGGGATTAGCTTGTTC
TACTATTACAACAATGGCTCAGCAGCCGCCAAAGTTTGCTGGAGTATCTTCTTTGTATA
GCTATGAGTATGACTATGGTAAGGGTAAATGGACTTGGTCTAATGAAGGCGGTTTCGA
TATTAAAGTGCCAGGGATTAAAATGAAGCCAAAAGAATGGATTTCTAAACAGGCTACT
TATCTTGAATTACAGCATTATATGCCTTATACTCCTGTTCTCGTGACATATGCTCATGAC
GTTCCTCCTAGCTCTATACTGTTATATCCGATGTCTGATCCTGATCAACTTGGAATAAA
TCGGCAGCAGCTGAAATTGAATTTGTATAGTTATTTTAACGATTTAAGACACGATTTTA
AATTAAAAGTTCTTGATGCACGTATTTCCAAAAATAAACAAAATATTGATACTATAAG
TAAATATTTACTAGAACTGGGTACTTATTTAGATGATTCTTATCGTATGATGGAACAAA
ATACACATAATATCAATAAAAATACACATAATATCAATAAGTTGTCTAAAGAATTGCA
AACTGGTTTAGCCAACCAATCAGCATTGTCTATGTTAGTGCAACCAAATGGTGTAGGC
AAAACGAGCGTTTCTGCTGCGGTAGGAGGTTATAGAGATAAAACTGCATTAGCCATTG
GTGTCGGCTCACGCATTACTGATCGCTTTACCGCTAAAGCGGGTGTAGCGTTCAATACC
TACAATGGCGGCATGTCTTATGGTGCTTCTGTTGGTTATGAATTCTAATCATTACGTTTA
ATCACTAATCGTTTTGGTTATAATAAAAAGGCTAAATGTTTCTCCTCACATTTAGCCTTT
CTTATTTATCTTTGTTATAGCTTTTGCTGTTATAAAACCGTTTTTTAGCCACTTTTATTAA
TTAAGCTTTTAAGCCTATTCAATCAGTTCTACTTTCACTTTTTTCACCATATTATCCGCC
ACTTCTAAAACGGTAATATTAAGTTGGTTTAGCCTAAATTGGGTACCTTCTATCGGAAT
TTTTTCTAAATGTTCTAAAATTAA

SEQ ID NO:4 PROTEIN SEQUENCE OF CIPA75 (set forth in FIG. 9)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLYSYEYDY GKGKWTWSNE
GGFDIKVPGI KMKPKEWISK QATYLELQHY MPYTPVLVTY AHDVPPSSIL
LYPMSDPDQL GINRQQLKLN LYSYFNDLRH DFKLKVLDAR ISKNKQNIDT
ISKYLLELGT YLDDSYRMME QNTHNINKNT HNINKLSKEL QTGLANQSAL
SMLVQPNGVG KTSVSAAVGG YRDKTALAIG VGSRITDRFT AKAGVAFNTY
NGGMSYGASV GYEF

SEQ ID NO:5 DNA SEQUENCE OF CIPA77

ATTTTATAATTTACAATACATTTTATATTTTTATATTATATAAATACGTCATTGACATTT
TTTTAAGGTAGAATAAGAAAGTAAATTCTATATTTACAATCAAGATTGACAATTATTTA
CTTAATGAGGTGATTATGAAAATTAAATGTTTAGTTGCCGTAGTGGGATTAGCTTGTTC
TACTATTACAACAATGGCTCAGCAGCCGCCAAAGTTTGCTGGAGTATCTTCTTTGTATA
GCTATGAGTATGACTATGGTAAGGGTAAATGGACTTGGTCTAATGAAGGCGGTTTCGA
TATTAAAGTGCCAGGGATTAAAATGAAGCCAAAAGAATGGATTTCTAAACAGGCTACT

TATCTTGAATTACAGCATTATATGCCTTATACTCCTGTTCTCGTGACATATGCTCATGAC
GTTCCTCCTAGCTCTATACTGTTATATCCGATGTCTGATCCTGATCAACTTGGAATAAA
TCGGCAGCAGCTGAAATTGAATTTGTATAGTTATTTTAACGATTTAAGACACGATTTTA
AATTAAAAGTTCTTGATGCACGTATTTCCAAAAATAAACAAAATATTGATACTATAAG
TAAATATTTACTAGAACTGGGTACTTATTTAGATGATTCTTATCGTATGATGGAACAAA
ATACACATAATATCAATAAAAATACACATAATATCAATAAGTTGTCTAAAGAATTGCA
AACTGGTTTAGCCAACCAATCAGCATTGTCTATGTTAGTGCAACCAAATGGTGTAGGC
AAAACGAGCGTTTCTGCTGCGGTAGGAGGTTATAGAGATAAAACTGCATTAGCCATTG
GTGTCGGCTCACGCATTACTGATCGCTTTACCGCTAAAGCGGGTGTAGCGTTCAATACC
TACAATGGCGGCATGTCTTATGGTGCTTCTGTTGGTTATGAATTCTAATCATTACGTTTA
ATCACTAATCG

SEQ ID NO:6 PROTEIN SEQUENCE OF CIPA77 (set forth in FIG. 9)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLYSYEYDY GKGKWTWSNE
GGFDIKVPGI KMKPKEWISK QATYLELQHY MPYTPVLVTY AHDVPPSSIL
LYPMSDPDQL GINRQQLKLN LYSYFNDLRH DFKLKVLDAR ISKNKQNIDT
ISKYLLELGT YLDDSYRMME QNTHNINKNT HNINKLSKEL QTGLANQSAL
SMLVQPNGVG KTSVSAAVGG YRDKTALAIG VGSRITDRFT AKAGVAFNTY
NGGMSYGASV GYEF

SEQ ID NO:7 DNA SEQUENCE OF CIP542 (Can)

TTTTATAATTTACAATACATTTTATATTTTTATATTATATAAATACGTCATTGACATTTTT
TTAATGTAAGGTAGAATAAGAAAGTAAATTCTATATTTACAATCAAGATTGACAATTA
TTTACTTAATGAGGTGATTATGAAAATTAAATGTTTAGTTGCCGTAGTGGGATTAGCTT
GTTCTACTATTACAACAATGGCTCAGCAGCCGCCAAAGTTTGCTGGAGTATCTTCTTTG
TATAGCTATGAGTATGACTATGGTAAGGGTAAATGGACTTGGTCTAATGAAGGCGGTT
TCGATATTAAAGTGCCAGGGATTAAAATGAAGCCAAAAGAATGGATTTCTAAACAGGC
TACTTATCTTGAATTACAGCATTATATGCCTTATACTCCTGTTCTCGTGACATATGCTCC
TGGCGTTTCTCCTAGCCCTATACTGTTATATCCGATGTCTGATCCTGATCAACTTGGAAT
AAATCGGCAGCAGCTGAAATTGAATTTGTATAGTTATTTTAACGATTTAAGACACGATT
TTAAATTAAAAGTTCTTGATGCACGTATTTCCAAAAATAAACAAAATATTGATACTATA
AGTAAATATTTACTAGAACTGGGTACTTATTTAGATGATTCTTATCGTATGATGGAACA
AAATACACATAATATCAATAAGTTGTCTAAAGAATTGCAAACTGGTTTAGCCAACCAA
TCAGCATTGTCTATGTTAGTGCAACCAAATGGTGTAGGCAAAACGAGCGTTTCTGCTGC
GGTAGGAGGTTATAGAGATAAAACTGCATTAGCCATTGGTGTCGGCTCACGCATTACT
GATCGCTTTACCGCTAAAGCGGGTGTAGCGTTCAATACCTTCTATCGGAATTTTTTCTA
AATGTTCTAAAATTA

SEQ ID NO:8 PROTEIN SEQUENCE OF CIP542 (Can) (set forth in FIG. 9)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLYSYEYDY GKGKWTWSNE
GGFDIKVPGI KMKPKEWISK QATYLELQHY MPYTPVLVTY APGVSPSPIL
LYPMSDPDQL GINRQQLKLN LYSYFNDLRH DFKLKVLDAR ISKNKQNIDT
ISKYLLELGT YLDDSYRMME QNTHNINKLS KELQTGLANQ SALSMLVQPN
GVGKTSVSAA VGGYRDKTAL AIGVGSRITD RFTAKAGVAF NT

SEQ ID NO:9 DNA SEQUENCE OF CIP542 (CDC)

AATGGCCATTTTATAATTTACAATACATTTTATATTTTTATATTATATAAATACGTCATT
GACATTTTTTTAATGTAAGGTAGAATAAGAAAGTAAATTCTATATTTACAATCAAGATT
GACAATTATTTACTTAATGAGGTGATTATGAAAATTAAATGTTTAGTTGCCGTAGTGGG
ATTAGCTTGTTCTACTATTACAACAATGGCTCAGCAGCCGCCAAAGTTTGCTGGAGTAT
CTTCTTTGTATAGCTATGAGTATGACTATGGTAAGGGTAAATGGACTTGGTCTAATGAA
GGCGGTTTCGATATTAAAGTGCCAGGGATTAAAATGAAGCCAAAAGAATGGATTTCTA
AACAGGCTACTTATCTTGAATTACAGCATTATATGCCTTATACTCCTGTTCTCGTGACA
TATGCTCCTGGCGTTTCTCCTAGCCCTATACTGTTATATCCGATGTCTGATCCTGATCAA
CTTGGAATAAATCGGCAGCAGCTGAAATTGAATTTGTATAGTTATTTTAACGATTTAAG
ACACGATTTTAAATTAAAAGTTCTTGATGCACGTATTTCCAAAAATAAACAAAATATTG
ATACTATAAGTAAATATTTACTAGAACTGGGTACTTATTTAGATGATTCTTATCGTATG
ATGGAACAAAATACACATAATATCAATAAGTTGTCTAAAGAATTGCAAACTGGTTTAG
CCAACCAATCAGCATTGTCTATGTTAGTGCAACCAAATGGTGTAGGCAAAACGAGCGT
TTCTGCTGCGGTAGGAGGTTATAGAGATAAAACTGCATTAGCCATTGGTGTCGGCTCA
CGCATTACTGATCGCTTTACCGCTAAAGCGGGTGTAGCGTTCAATACCTACAATGGCG
GCATGTCTTATGGTGCTTCTGTTGGTTATGAATTCTAATCATTACGTTTAATCACTAATC
GTTTTGGTTATAATAAAAAGGCTAAATGTTTCTCCTCACATTTAGCCTTTCTTATTTATC
TTTGTTATAGCCTTTTGCTGTTATAAAACCGTTTTTTAGCCACTTTTATTAATTAAGCTTT
TAAGCCTATTCAATCAGTTCTACTTTCACTTTTTTCACCATATTATCCGCCACTTCTAAA
ACGGTAATATTAAGTTGGTTTAGCCTAAATTGGGTACCTTCTATCGGAATTTTTTCTAA
ATGTTCTAAAATTAA

SEQ ID NO:10 PROTEIN SEQUENCE OF CIP542 (CDC) (set forth in FIG. 9)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLYSYEYDY GKGKWTWSNE
GGFDIKVPGI KMKPKEWISK QATYLELQHY MPYTPVLVTY APGVSPSPIL
LYPMSDPDQL GINRQQLKLN LYSYFNDLRH DFKLKVLDAR ISKNKQNIDT
ISKYLLELGT YLDDSYRMME QNTHNINKLS KELQTGLANQ SALSMLVQPN
GVGKTSVSAA VGGYRDKTAL AIGVGSRITD RFTAKAGVAF NTYNGGMSYG
ASVGYEF

SEQ ID NO:11 DNA SEQUENCE OF CHIA

TATTTACAATCAAGATTGACAATTATTTACTTAATGAGGTGATTATGAAAATTAAATGT
TTAGTTGCCGTAGTGGGATTAGCTTGTTCTACTATTACAACAATGGCTCAGCAGCCGCC
AAAGTTTGCTGGAGTATCTTCTTTGGATAGCTATGAGTATGACTATGGTAAGGGTAAAT
GGACTTGGTCTGAAAAAGACGGTTTCGATATTAAAGCGCCAGGGATTAAAATGAAGCC
AAAAAAATGGATTTCTAGACAGGCTACTTATCTTGGATTACAGCATTATATGCCTTATA
CTCCTGTTCTCGTGACATATGCTTCTGCAGAACCTAACACTGTACTGTTATATCCGATG
CCTGATCCTGATCAACTTGGAATAAATCGGCAGCAGCTGAAATTGAATTTGTATAGTTA
TTTTAACGATTTAAGACACGGTTTTAAATTAAATGTTCTTGATGCACGTATTTCCCAAA
ATAAACAAAATATTGATACTATAAGTGAATATTTACTAAAACTGGGTACTTATTTAGAT
AGTTCTTATCGTATGATGGAACAAAATACACATAATATCAATAAAAATACACATAATA
TCAATAAGTTGTCTAAAGAATTGCAAACTGGTTTAGCCAACCAATCAGCATTGTCTATG
TTAGTGCAACCAAATGGTGTAGGCAAAACGAGCGTTTCTGCTGCGGTAGGAGGTTATA
GAGATAAAACTGCATTAGCCATTGGTGTCGGCTCACGCATTACTGATCGCTTTACCGCT
AAAGCGGGTGTAGCGTTCAATACCTACAATGGCGGCATGTCTTATGGTGCTTCTGTTGG
TTATGAATTCTAATCATTACGTTTAATCACTAATCGTTTTGGTTATAATAAAAAGGCTA
AATGTTTCTCCTCACATTTAGCCTTTCTTATTTATCTTTGT

SEQ ID NO:12 PROTEIN SEQUENCE OF CHIA(set forth in FIG. 9)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLDSYEYDY GKGKWTWSEK
DGFDIKAPGI KMKPKKWISR QATYLGLQHY MPYTPVLVTY ASAEPNTVLL
YPMPDPDQLG INRQQLKLNL YSYFNDLRHG FKLNVLDARI SQNKQNIDTI
SEYLLKLGTY LDSSYRMMEQ NTHNINKNTH NINKLSKELQ TGLANQSALS
MLVQPNGVGK TSVSAAVGGY RDKTALAIGV GSRITDRFTA KAGVAFNTYN
GGMSYGASVG YEF

SEQ ID NO:13 DNA SEQUENCE OF V-1157

CTTTTATAATTTACAATACATTTTATATTTTTATATTATATAAATACGTCATTGACATTT
TTTTAATGTAAGGTAGAATAAGAAAGTAAATTCTATATTTACAATCAAGATTGACAATT
ATTTACTTAATGAGGTGATTATGAAAATTAAATGTTTAGTTGCCGTAGTGGGATTAGCT
TGTTCTACTATTACAACAATGGCTCAGCAGCCGCCAAAGTTTGCTGGAGTATCTTCTTT
GTATAGCTATGAGTATGACTATGGTAAGGGTAAATGGACTTGGTCTAATGAAGGCGGT
TTCGATATTAAAGTGCCAGGGATTAAAATGAAGCCAAAAGAATGGATTTCTAAACAGG
CTACTTATCTTGAATTACAGCATTATATGCCTTATACTCCTGTTCTCGTGACATCTGCTC
CTGACGTTCCTCCTAGCTCTATACTGTTATATCCGATGTCTGATCCTGATCAACTTGGA
ATAAATCGGCAGCAGCTGAAATTGAATTTGTATAGTTATTTTAACGATTTAAGACACG
ATTTTAAATTAAAAGTTCTTGATGCACGTATTTCCAAAAATAAACAAAATATTGATACT

ATAAGTAAATATTTACTAGAACTGGGTACTTATTTAGATGGTTCTTATCGTATGATGGA
ACAAAATACACATAATATCAATAAAAATACACATAATATCAATAAAAATACACATAAT
ATCAATAAGTTGTCTAAAGAATTGCAAACTGGTTTAGCCAACCAATCAGCATTGTCTAT
GTTAGTGCAACCAAATGGTGTAGGCAAAACGAGCGTTTCTGCTGCGGTAGGAGGTTAT
AGAGATAAAACTGCATTAGCCATTGGTGTCGGCTCACGCATTACTGATCGCTTTACCGC
TAAAGCGGGTGTAGCGTTCAATACCTACAATGGCGGCATGTCTTATGGTGCTTCTGTTG
GTTATGAATTCTAATCATTACGTTTAATCACTAATCGTTTTGGTTATAATAAAAAGGCT
AAATGTTTCTCCTCACATTTAGCCTTTCTTATTTATCTTTGTTATAGCTTTTGCTGTTATA
AAACCGTTTTTTAGCCACTTTTATTAATTAAGCTTTTAAGCCTATTCAATCAGTTCTACT
TTCACTTTTTTCACCATATTATCCGCCACTTCTAAAACGGTAATATTAAGTTGGTTTAGC
CTAAATTGGGTACCTTCTATCGGAATTTTTTCTAAATGTTCTAAAATTAA

SEQ ID NO:14 PROTEIN SEQUENCE OF V-1157 (set forth in FIG. 9)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLYSYEYDY GKGKWTWSNE
GGFDIKVPGI KMKPKEWISK QATYLELQHY MPYTPVLVTS APDVPPSSIL
LYPMSDPDQL GINRQQLKLN LYSYFNDLRH DFKLKVLDAR ISKNKQNIDT
ISKYLLELGT YLDGSYRMME QNTHNINKNT HNINKNTHNI NKLSKELQTG
LANQSALSML VQPNGVGKTS VSAAVGGYRD KTALAIGVGS RITDRFTAKA
GVAFNTYNGG MSYGASVGYE F

SEQ ID NO:15 DNA SEQUENCE OF M90-02

TTTTATAATTTACAATACATTTTATATTTTTATATTATATAAATACCGTCATTGACATTT
TTTTAATGTAAGGTAGAATAAGAAAGTAAATTCTATATTTACAATCAAGATTGACAATT
ATTTACTTAATGAGGTGATTATGAAAATTAAATGTTTAGTTGCCGTAGTGGGATTAGCT
TGTTCTACTATTACAACAATGGCTCAGCAGCCGCCAAAGTTTGCTGGAGTATCTTCTTT
GTATAGCTATGAGTATGACTATGGTAAGGGTAAATGGACTTGGTCTAATGAAGGCGGT
TTCGATATTAAAGTGCCAGGGATTAAAATGAAGCCAAAAGAATGGATTTCTAAACAGG
CTACTTATCTTGAATTACAGCATTATATGCCTTATACTCCTGTTCTCGTGACATCTGCTC
CTGACGTTTCTCCTAGCTCTATCTCTATACTGTTATATCCGATGTCTGATCCTGATCAAC
TTGGAATAAATCGGCAGCAGCTGAAATTGAATTTGTATAGTTATTTTAACGATTTAAGA
CACGATTTTAAATTAAAAGTTCTTGATGCACGTATTTCCAAAAATAAACAAAATATTGA
TACTATAAGTAAATATTTACTAGAACTGGGTACTTATTTAGATGGTTCTTATCGTATGA
TGGAACAAAATACACATAATATCAATAAAAATACACATAATATCAATAAAAATACACA
TAATATCAATAAGTTGTCTAAAGAATTGCAAACTGGTTTAGCCAACCAATCAGCATTGT
CTATGTTAGTGCAACCAAATGGTGTAGGCAAAACGAGCGTTTCTGCTGCGGTAGGAGG
TTATAGAGATAAAACTGCATTAGCCATTGGTGTCGGCTCACGCATTACTGATCGCTTTA
CCGCTAAAGCGGGTGTAGCGTTCAATACCTACAATGGCGGCATGTCTTATGGTGCTTCT

GTTGGTTATGAATTCTAATCATTACGTTTAATCACTAATCGTTTTGGTTATAATAAAAA
GGCTAAATGTTTCTCCTCACATTTAGCCTTTTCTTATTTATCTTT

SEQ ID NO:16 PROTEIN SEQUENCE OF M90-02 (set forth in FIG. 9)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLYSYEYDY GKGKWTWSNE
GGFDIKVPGI KMKPKEWISK QATYLELQHY MPYTPVLVTS APDVSPSSIS
ILLYPMSDPD QLGINRQQLK LNLYSYFNDL RHDFKLKVLD ARISKNKQNI
DTISKYLLEL GTYLDGSYRM MEQNTHNINK NTHNINKNTH NINKLSKELQ
TGLANQSALS MLVQPNGVGK TSVSAAVGGY RDKTALAIGV GSRITDRFTA
KAGVAFNTYN GGMSYGASVG YEF

SEQ ID NO:17 DNA SEQUENCE OF 406

ATTTTATAATTTACAATACATTTTTATTTTTATATTATATAAATACGTCATTGACATTTTT
TTAATGTAAGGTAGAATAAGAAAGTAAATTCTATATTTACAATCAAGATTGACAATTA
TTTACTTAATGAGGTGATTATGAAAATTAAATGTTTAGTTGCCGTAGTGGGATTAGCTT
GTTCTACTATTACAACAATGGCTCAGCAGCCGCCAAAGTTTGCTGGAGTATCTTCTTTG
TATAGCTATGAGTATGACTATGGTAAGGGTAAATGGACTTGGTCTAATGAAGGCGGTT
TCGATATTAAAGTGCCAGGGATTAAAATGAAGCCAAAAGAATGGATTTCTAAACAGGC
TACTTATCTTGAATTACAGCATTATATGCCTTATACTCCTGTTCTCGTGACATATGCTCC
TGGCGTTTCTCCTAGCCCTATACTGTTATATCCGATGTCTGATCCTGATCAACTTGGAAT
AAATCGGCAGCAGCTGAAATTGAATTTGTATAGTTATTTTAACGATTTAAGACACGATT
TTAAATTAAAAGTTCTTGATGCACGTATTTCCAAAAATAAACAAAATATTGATACTATA
AGTAAATATTTACTAGAACTGGGTACTTATTTAGATGATTCTTATCGTATGATGGAACA
AAATACACATAATATCAATAAGTTGTCTAAAGAATTGCAAACTGGTTTAGCCAACCAA
TCAGCATTGTCTATGTTAGTGCAACCAAATGGTGTAGGCAAAACGAGCGTTTCTGCTGC
GGTAGGAGGTTATAGAGATAAAACTGCATTAGCCATTGGTGTCGGCTCACGCATTACT
GATCGCTTTACCGCTAAAGCGGGTGTAGCGTTCAATACCTACAATGGCGGCATGTCTTA
TGGTGCTTCTGTTGGTTATGAATTCTAATCATTACGTTTAATCACTAATCGTTTTGGTTA
TAATAAAAAGGCTAAATGTTTCTCCTCACATTTAGCCTTTCTTATTTATCTTTGTTATAG
CTTTTGCTGTTATAAAACCGTTTTTTAGCCACTTTTATTAATTAAGCTTTTAAGCCTATT
CAATCAGTTCTACTTTCACTTTTTTCACCATATTATCCGCCACTTCTAAAACGGTAATAT
TAAGTTGGTTTAGCCTAAATTGGGTACCTTCTATCGGAATTTTTTCTAAATGTTCTAAA
ATTAAG

SEQ ID NO:18 PROTEIN SEQUENCE OF 406 (set forth in FIG. 9)

MKIKCLVAVV GLACSTITTM AQQPPKFAGV SSLYSYEYDY GKGKWTWSNE
GGFDIKVPGI KMKPKEWISK QATYLELQHY MPYTPVLVTY APGVSPSPIL

LYPMSDPDQL GINRQQLKLN LYSYFNDLRH DFKLKVLDAR ISKNKQNIDT
ISKYLLELGT YLDDSYRMME QNTHNINKLS KELQTGLANQ SALSMLVQPN
GVGKTSVSAA VGGYRDKTAL AIGVGSRITD RFTAKAGVAF NTYNGGMSYG
ASVGYEF